

```

1  CGGGCAGCAA AGGAGGATGG CGAGGGGCTG ATACTGAACC CGGGAAGGGT
51  GGGCTGTGCT GAAGCTAGAG CCGGAGCCGG AGCTGGGGCC AGAACC CGAG
101  CACTGCCATG TCCACGCAGA GACTTCGGAA TGAAGACTAC CACGACTACA
151  GCTCCACGGA CGTGAGCCCT GAGGAGAGCC CGTCGGAAGG CCTCAACAAC
201  CTCTCTCTCC CGGGCTCCTA CCAGCGCTTT GGTCAAAGCA ATAGACAAC
251  ATGGTCCAGC ACCTTGATCC ACCTGTAAAA AGGCAACATT GGACAGGAC
301  TCCTGGGATG CCCTCTGGCG GTGAAAAATG CAGGCATCGT GATGGGTCCC
351  ATCAGCCTGC TGATCATAGG CATCGTGCC GTGCACTGCA TGGGTATCCT
401  GGTGAAATGT GCTCACCCT TCTGCCGCG GCTGAATAAA TCCTTTGTGG
451  ATTATGGTGA TACTGTGATG TATGGACTAG AATCCAGCCC CTGCTCCTGG
501  CTCCGGAACC ACGCACACTG GGAAGACGT GTTGTGGACT TCTTCCTGAT
551  TGTCAACCAG CTGGGATTC TCTGTGTCTA TTTTGTGTTT CTGGCTGACA
601  ACTTTAAACA GGTGATAGAA GCGGCCAATG GGACCACCAA TAACTGCCAC
651  AACAAATGAGA CGGTGATTCT GACGCCTACC ATGGACTCGC GACTCTACAT
701  GCTCTCCTTC CTGCCCTTCC TGGTGCTGCT GGTTTTCATC AGGAACCTCC
751  GAGCCCTGTC CATCTTCTCC CTGTTGGCCA ACATCACTAT GCTGGTCAGC
801  TTGGTCATGA TCTACCACTT CATTTGTCAG AGGATCCAG ACCCCAGCCA
851  CCTCCCTTGG TGGGCCCTT GGAAGACCTA CCCTCTCTTC TTTGGCACAG
901  CGATTTTTTC ATTTGAAGGC ATTGGAATGG TTCTGCCCCT GGAAAAAACA
951  ATGAAGGATG CTCGGAAGTT CCCACTCATC CTGTACCTGG GCATGGTCAT
1001  CGTCACCATC CTCTACATCA GCCTGGGGTG TCTGGGGTAC CTGCAATTTG
1051  GAGCTAATAT CCAAGGCAGC ATAACCCTCA ACCTGCCCAA CTGCTGGTTG
1101  TACCAGTCAG TTAAGCTGCT GACTCCATC GGGATCTTTT TCACCTACGC
1151  ACTCCAGTTC TACGTCCC GG CTGAGATCAT CATCCCCTTC TTTGTGTCCC
1201  GAGCGCCCGA GCACTGTGAG TTAGTGGTGG ACCTGTTTGT GCGCACAGTG
1251  CTGGTCTGCC TGACATGCAT CTTGGCCATC CTCATCCCCC GCCTGGACCT
1301  GGTCACTCTC CTGGTGGGCT CCGTGAGCAG CAGCGCCCTG GCCCTCATCA
1351  TCCCACCGCT CTGGAGGTC ACCACCTTCT ACTCAGAGGG CATGAGCCCC
1401  CTCAACATCT TTAAGGACGC CCTGATCAGC ATCCTGGGCT TCGTGGGCTT
1451  TGTGGTGGGG ACCTATGAGG CTCTCTATGA GCTGATCCAG CCAAGCAATG
1501  CTCCCATCTT CATCAATTCC ACCTGTGCCT TCATATAGGG ATCTGGGTTC
1551  GTCTCTGCAG CTGCCATACC CTGCCCATG TGTCCTCCCT TACCTGTCTC
1601  CAGAGCCTCA GGTATGGTCC AGGCTCTGAG GAAAGTCAGG GTTGCTGTGT
1651  GGGAAACCCT CTGCCCTGCA CTGGGATACC CTGGGCCAGG TAACCTGAGG
1701  GCAGGGGAGA GGTGGGGTGG CAGACAGGCA GAAGTGCTAC TAGTGACAGG
1751  GCTGCCATCG CTCACCTGTA CCTATTACA CCCAGAACTT TCCAGCTCCC
1801  CCTCATCATG CCTCCTCCTT CTAACCTGCC TCCCCTCTGC TGGTGCACCT
1851  CGCCCAACTC ATTCTTACTG CACAGTTCAC TTTATTTAAC AATTTTCATG
1901  TCCCCCATCT CGCTCTGTGC CCCTCCCAC CAGGGCTTCA GCAGGAGCCC
1951  TGGACTCATC ATCAATAAAC ACTGTTACAG CAAAAA AAAAAA
2001  AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA
2051  AAAAAA AAAAAA AAAAAA AAAAAA AAA (SEQ ID NO:1)

```

FEATURES:
 5'UTR: 1-107
 Start Codon: 108
 Stop Codon: 1536
 3'UTR: 1539

FIGURE 1A

HOMOLOGOUS PROTEINS:

Top BLAST Hits:

Top BLAST Hits:				Score	E
CRA 89000000199482	/altid=gi 7297404	/def=gb AAF52663.1	(AE003...	330	4e-89
CRA 89000000199480	/altid=gi 7297402	/def=gb AAF52661.1	(AE003...	330	4e-89
CRA 89000000199481	/altid=gi 7297403	/def=gb AAF52662.1	(AE003...	330	4e-89
CRA 89000000197173	/altid=gi 7294781	/def=gb AAF50116.1	(AE003...	268	1e-70
CRA 89000000195851	/altid=gi 7293314	/def=gb AAF48694.1	(AE003...	265	1e-69
CRA 18000005127815	/altid=gi 7509795	/def=pir T26845	hypotheti...	258	2e-67
CRA 89000000194855	/altid=gi 7292192	/def=gb AAF47603.1	(AE003...	253	5e-66
CRA 89000000197171	/altid=gi 7294779	/def=gb AAF50114.1	(AE003...	252	8e-66
CRA 89000000197172	/altid=gi 7294780	/def=gb AAF50115.1	(AE003...	252	8e-66
CRA 18000005102492	/altid=gi 2429516	/def=gb AAB71045.1	(AF025...	250	3e-65

BLAST dbEST hits:

	Score	E
gi 5422591 /dataset=dbest /taxon=9606 ...	1400	0.0
gi 3648072 /dataset=dbest /taxon=9606 ...	1017	0.0
gi 5746200 /dataset=dbest /taxon=9606 ...	730	0.0
gi 10249244 /dataset=dbest /taxon=96...	642	0.0
gi 8612353 /dataset=dbest /taxon=960...	329	1e-87
gi 10083945 /dataset=dbest /taxon=960...	313	7e-83

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

From BLAST dbEST hits:

gi|5422591 testis
 gi|3648072 Testis
 gi|5746200 Brain meningiomas
 gi|10249244 Brain normal
 gi|8612353 Head-neck
 gi|10083945 colon

From tissue screening panels:

Human whole liver

1 MSTQRLRNED YHDYSSTDVS PEESPSEGLN NLSSPGSYQR FGQSNSTTW
 51 QTLIHLLKGN IGTGLLGLPL AVKNAGIVMG PISLLIIGIV AVHCMGILVK
 101 CAHHFCRRLN KSFVDYGDV MYGLESPCS WLRNHAHWGR RVVDFFLIVT
 151 QLGFCVCYFV FLADNFKQVI EAANGTTNNC HNNETVILT TMSRLYMLS
 201 FLPLVLLVF IRNLRLSIF SLLANITMLV SLVMIYQFIV QRIPDP SHLP
 251 LVAPWKTYPL FFGTAIFSFE GIGMVLPLEN KMKDPRKFPL ILYLGMVIVT
 301 ILYISLGCLG YLQFGANIQG SITLNLPCW LYQSVKLLYS IGIFFTYALQ
 351 FYVPAEIIIP FFVSRAPHC ELVVDLFVRT VLVCLTCILA ILIPRLDLVI
 401 SLVGSVSSSA LALIIPLLE VTTFYSEGMS PLTIFKDALI SILGFVGFVV
 451 GTYEALYELI QPSNAIFIN STCAFI (SEQ ID NO:2)

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
 N-glycosylation site

Number of matches: 7

1	31-34	NLSS
2	45-48	NSTT
3	110-113	NKSF
4	174-177	NGTT
5	183-186	NETV
6	225-228	NITM
7	470-473	NSTC

[2] PDOC00005 PS00005 PKC_PHOSPHO_SITE
 Protein kinase C phosphorylation site

Number of matches: 2

1	3-5	TQR
2	334-336	SVK

[3] PDOC00006 PS00006 CK2_PHOSPHO_SITE
 Casein kinase II phosphorylation site

Number of matches: 4

1	15-18	SSTD
2	20-23	SPEE
3	24-27	SPSE
4	112-115	SFVD

[4] PDOC00007 PS00007 TYR_PHOSPHO_SITE
 Tyrosine kinase phosphorylation site

7-14 RNEDYHDY

[5] PDOC00008 PS00008 MYRISTYL
 N-myristoylation site

Number of matches: 7

1	42-47	GQSNST
2	59-64	GNIGTG
3	67-72	GLPLAV
4	175-180	GTTNNC
5	342-347	GIFTTY
6	404-409	GSVSSS
7	451-456	GTYEAL

[6] PDOC00009 PS00009 AMIDATION
 Amidation site

138-141 WGRR

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	52	72	0.668	Putative
2	75	95	2.032	Certain
3	143	163	1.799	Certain

FIGURE 2A

Docket No.: CL001062CON
Serial No.: TO BE ASSIGNED
Inventors: WEI, Ming-Hui et al.
Title: ISOLATED HUMAN TRANSPORTER PROTEINS...

4	193	213	1.467	Certain
5	216	236	1.884	Certain
6	258	278	1.566	Certain
7	289	309	2.126	Certain
8	335	355	1.378	Certain
9	375	395	1.332	Certain
10	398	418	1.748	Certain
11	437	457	1.533	Certain

FIGURE 2B

BLAST Alignment to Top Hit:

>CRA|89000000199482 /altid=gi|7297404 /def=gb|AAF52663.1| (AE003621)
 CG13384 gene product [alt 3] [Drosophila melanogaster]
 /org=Drosophila melanogaster /taxon=7227 /dataset=nraa
 /length=486
 Length = 486

Score = 330 bits (837), Expect = 2e-89
 Identities = 184/425 (43%), Positives = 262/425 (61%), Gaps = 32/425 (7%)

Query: 47 TTWFQTLIHLLKGNIGTGLLGLPLAVKNAGIVMGPIISLLIIGIVAVHCGMILVKCAHHFC 106
 T+ F TL+HLLKGNIGTG+L +P A KNAG+ +G +I+G+ HCM +LV C+H C
 Sbjct: 78 TSNFDTLVHLLKGNIGTGILAMPDAFNAGLYVGLFGTMIMGAICTHGMHMLVNCSELHC 137

Query: 107 RRLNKSFDVYGDVTMYGLESSPCSWLRNHAHWGRRVVDFFLIVTQLGFCCVYFVFLADNF 166
 RR + +D+ + ES P LR ++ RR+V FL +TQ+GFCCVYF+F+A N
 Sbjct: 138 RRFQPSLDFSEVAYCSFESGPLG-LRRYSMLARRIVTTFITQIGFCCVYFLFVALNI 196

Query: 167 KQVIEAANGTTNNCHNETVILTPTMDSRLYMLSFLPFLVLLVFIRNLRALSIFSLLANI 226
 K V++ H + M ++Y+L L ++LL +RNL+ L+ SL+A +
 Sbjct: 197 KDVM-----HYK-----MPVQIYLLIMLGPILLNLVRNLKYLTPVSLVAAL 240

Query: 227 TMLVSLVMIYQFIVQRIIPDPSHLPLVAPWKTYPLFFGTAFSFEIGMVLPLENKMKDPR 286
 + L + + + + +PD + VA W T PL+FGTAI++FEGIG+VLPLEN M+ P
 Sbjct: 241 LTVAGLAITFSYMLVDLPDVHTVKPVATWATLPLYFGTAIYAFEGIGVVLPLENNMRTPE 300

Query: 287 K----PLILYLMGVIVTILYISLGCLGYLQFGANIQGSITLNLN-NCWLQYSVKLLYSIG 342
 F +L GMVIV LY ++G GYL++G +++GSITLNL L Q V++ ++
 Sbjct: 301 DFGGTTGLVNTGMVIVACLYTAVGFFGYLYKGEHVEGSITLNLPGQDTSQLVRISMAVA 360

Query: 343 IFFTYALQFYVPAEIIIPF-----VSRAPHECELVDLFVRTVLVCLTCILAILIPRLD 397
 IF +Y LQFYVP I+ PF +RA+ V +R VLV T +LA IP L
 Sbjct: 361 IFLSYTLQFYVPVNIPEFVRSHFDTTRAKDLSATV----LRVVLVTFTFLATCIPNLG 416

Query: 398 LVISLVGSVSSSALALIIPPLLEVTTFYSEGMSPLT--IFKDALISILGFVGFVGTYE 455
 +ISLVG+VSSSALALI PP++EV TFY+ G ++KD LI I G GFV GT+ +
 Sbjct: 417 SIISLVGAVSSSALAI TAPPIIEVITFYNVGYGRFNWMLWKDVLILIFGLCGFVFGTWA 476

Query: 456 LYELI 460
 L +++
 Sbjct: 477 LAQIL 481 (SEQ ID NO:4)

>CRA|89000000199481 /altid=gi|7297403 /def=gb|AAF52662.1| (AE003621)
 CG13384 gene product [alt 2] [Drosophila melanogaster]
 /org=Drosophila melanogaster /taxon=7227 /dataset=nraa
 /length=482
 Length = 482

Score = 330 bits (837), Expect = 2e-89
 Identities = 184/425 (43%), Positives = 262/425 (61%), Gaps = 32/425 (7%)

Query: 47 TTWFQTLIHLLKGNIGTGLLGLPLAVKNAGIVMGPIISLLIIGIVAVHCGMILVKCAHHFC 106
 T+ F TL+HLLKGNIGTG+L +P A KNAG+ +G +I+G+ HCM +LV C+H C
 Sbjct: 74 TSNFDTLVHLLKGNIGTGILAMPDAFNAGLYVGLFGTMIMGAICTHGMHMLVNCSELHC 133

Query: 107 RRLNKSFDVYGDVTMYGLESSPCSWLRNHAHWGRRVVDFFLIVTQLGFCCVYFVFLADNF 166
 RR + +D+ + ES P LR ++ RR+V FL +TQ+GFCCVYF+F+A N
 Sbjct: 134 RRFQPSLDFSEVAYCSFESGPLG-LRRYSMLARRIVTTFITQIGFCCVYFLFVALNI 192

Query: 167 KQVIEAANGTTNNCHNETVILTPTMDSRLYMLSFLPFLVLLVFIRNLRALSIFSLLANI 226
 K V++ H + M ++Y+L L ++LL +RNL+ L+ SL+A +
 Sbjct: 193 KDVM-----HYK-----MPVQIYLLIMLGPILLNLVRNLKYLTPVSLVAAL 236

Query: 227 TMLVSLVMIYQFIVQRIIPDPSHLPLVAPWKTYPLFFGTAFSFEIGMVLPLENKMKDPR 286
 + L + + + + +PD + VA W T PL+FGTAI++FEGIG+VLPLEN M+ P
 Sbjct: 237 LTVAGLAITFSYMLVDLPDVHTVKPVATWATLPLYFGTAIYAFEGIGVVLPLENNMRTPE 296

FIGURE 2C

Query: 287 KF---PLILYLGMIIVITILYISLGLGYLQFGANIQGSITLNLPCWLYQSVKLLYSIG 342
 F +L GMVIV LY ++G GYL++G +++GSITLNL L Q V++ ++
 Sbjct: 297 DFGGTTGVLNTGMVIVACLYTAVGFFGYLKYGEHVEGSITLNLPGDQLSGLVRISMAVA 356

Query: 343 IFFTYALQFYVPAEIIIPFF-----VSRAPHECELVDLFRVTVLVCLTCILAILIPRLD 397
 IF +Y LQFYVP I+ PF +RA + V +R VLV T +LA IP L
 Sbjct: 357 IFLSYTLQFYVPVNIPEPFVRSHFDITRAKDL SATV----LRVVLVTFTLLATCIPNLG 412

Query: 398 LVISLVGVSSSALALIIPPLLEVTTFYSEGMSPLT--IFKDALISILGFVGFVVGTYEA 455
 +ISLVG+VSSSALALI PP++EV TFY+ G ++KD LI I G GFV GT+ +
 Sbjct: 413 SIISLVGAVSSSALALIAPPIIEVITFYNVGYGRFNWMLWKDVLILIFGLCGFVFGTWAS 472

Query: 456 LYELI 460
 L +++
 Sbjct: 473 LAQIL 477 (SEQ ID NO:5)

>CRA|18000005127815 /altid=gi|7509795 /def=pir||T26845 hypothetical
 protein Y43F4B.7 - Caenorhabditis elegans
 /org=Caenorhabditis elegans /taxon=6239 /dataset=nraa
 /length=607
 Length = 607

Score = 258 bits (652), Expect = 1e-67
 Identities = 142/418 (33%), Positives = 235/418 (55%), Gaps = 19/418 (4%)

Query: 40 RFGQSNSTTWFTQLIHLLKGNIGTGLLGLPLAVKNAGIVMGPISSLIIIGIVAVHCMGILV 99
 R NS T Q IH++K +GTGLL LPLA K++G+ +G I ++I ++ ++CM +V
 Sbjct: 42 RLPTENSLTPEQAFIHMVKAMLTGGLSLPLAFKHSGFLGLILTVLICLICLYCMRQVV 101

Query: 100 KCAHFCRRLNKSFDVYGDVTMYGLESPCSWLNRNHAHWGRRVDFFLIVTQLGFCCVYF 159
 AH C R + +DY + + +E P W++ + ++ ++V+ + ++QLGFCCVYF
 Sbjct: 102 FAAHFVCRNRGRDLIDYANIMRGAVEMGP-PWIKRNGYFFKQLVNVNMFISQLGFCCVYF 160

Query: 160 VFLADNFKQVIEAANGTTNNCHNETVILTPTMDSRLYMLSFLPFLVLLVFIRNLRLALS 219
 VF+ADN + NN T I + ++ML L ++ + IR L L+
 Sbjct: 161 VFMADNLEDF-----NNNTSI---HLSKAVWMLLLLIPMLISICSIRRLSLAP 206

Query: 220 FSLANITMLVSLVMIYQFIVQRIIPDPSHLPLVAPWKTYPLFFGTAFSFEIGIMVLPLE 279
 F++ AN+ +V++ ++ F + + S LP PLFFGT +F+FE+ ++P+E
 Sbjct: 207 FAMAANVVVVAVAVVLFFLSDLRPISSLPWFGKATDLPLFFGTVMFAFEGVAVIMPIE 266

Query: 280 NKMKDPRKFL---ILYLGMIIVITILYISLGLGYLQFGANIQGSITLNLPCWLYQSVK 336
 N+m+ P F +L ++V ++ G GYL G +++ +TLNLP YQ++K
 Sbjct: 267 NRMQSPHAFISWGNLSSCLVLAIFSVTFYGYLSLGNVVDKDTATLNLPMTPFYQTIK 326

Query: 337 LLYSIGIFFTYALQFYVPAEIIIPFFVSRAPHECELVDLFRVTVLVCLTCILAILIPRL 396
 L++ I +Y LQFYVP E I + + P ++ + R V LTC +A LIP L
 Sbjct: 327 LMFVACIMISYPLQFYVPMERIEKWITRKIPVDKQTLIYIARYSGVILTCAIAELIPHL 386

Query: 397 DLVISLVGVSSSALALIIPPLLEVTTFYSEG-MSPLTIFKDALISILGFVGFVVGTY 453
 L ISL+G+ S +++AL+ PP +E+ T Y++ +S K+ ++ F+GF GTY
 Sbjct: 387 ALFISLIGAFSGASMALLFPPCIELLTSYAKNELSTGLWIKNIVLLTFAFIFGTTGT 444 (SEQ
 ID NO:6)

>CRA|335001101719045 /dataset=FastAlert /length=476
 /altid=Derwent|wo200071709.21
 Length = 476
 Score = 909 bits (2324), Expect = 0.0
 Identities = 450/476 (94%), Positives = 459/476 (95%)
 Frame = +3

Query: 108 MSTQRLRNEDYHDYSSDVSPEESPSEGLNNLSSPGSYQRFQGSNSTTWFTQLIHLLKGN 287
 MSTQRLRNEDYHDYSSDVSPEESPSEGLNNLSSPGSYQRFQGSNSTTWFTQLIHLLKGN
 Sbjct: 1 MSTQRLRNEDYHDYSSDVSPEESPSEGLNNLSSPGSYQRFQGSNSTTWFTQLIHLLKGN 60

FIGURE 2D

Query: 288 IGTGLGLPLAVKNAGIVMGPIISLLIIGIVAVHCMGILVKCAHHFCRRLNKSFVDYGDV 467
IGTGLGLPLAVKNAGIVMGPIISLLIIGIVAVHCMGILVKCAHHFCRRLNKSFVDYGDV
Sbjct: 61 IGTGLGLPLAVKNAGIVMGPIISLLIIGIVAVHCMGILVKCAHHFCRRLNKSFVDYGDV 120

Query: 468 MYGLESPCSWLRNHAHWGRRVDFFLIVTQLGFCCVYFVFLADNFKQVIEAANGTTNNC 647
MYGLESPCSWLRNHAHWGRRVDFFLIVTQLGFCCVYFVFLADNFKQVIEAANGTTNNC
Sbjct: 121 MYGLESPCSWLRNHAHWGRRVDFFLIVTQLGFCCVYFVFLADNFKQVIEAANGTTNNC 180

Query: 648 HNNETVILTPTMDSRLYMLSFLPFLVLLVFIRNLRALSIFSLLANITMLVSLVMIYQFIV 827
HNNETVILTPTMDSRLYMLSFLPFLVLLVFIRNLRALSIFSLLANITMLVSLVMIYQFIV
Sbjct: 181 HNNETVILTPTMDSRLYMLSFLPFLVLLVFIRNLRALSIFSLLANITMLVSLVMIYQFIV 240

Query: 828 QRIPDPSHLPLVAPWKTYPLFFGTAFISFEGIGMVLPLENKMKDPRKFPLILYLGMIIVT 1007
QRIPDPSHLPLVAPWKTYPLFFGTAFISFEGIGMVLPLENKMKDPRKFPLILYLGMIIVT
Sbjct: 241 QRIPDPSHLPLVAPWKTYPLFFGTAFISFEGIGMVLPLENKMKDPRKFPLILYLGMIIVT 300

Query: 1008 ILYISLGCLGYLQFGANIQGSITLNLPCWLYQSVKLLYSIGIFFTYALQFYVPAEIIIP 1187
ILYISLGCLGYLQFGANIQGSITLNLPCWLYQSV+LLY GI TY LQFYV A+II+P
Sbjct: 301 ILYISLGCLGYLQFGANIQGSITLNLPCWLYQSV+LLY LGGICLTLYPLQFYVSAKIIIP 360

Query: 1188 FFVSRAPHEHCELVDLFRVTVLVCLTCILAILIPRLDLVISLVGSVSSSALALIIPPLLE 1367
VS + C L+VDL + + ++C TCILAILIPRLDLVISLVGSVSSSALALIIPPLLE
Sbjct: 361 VIVSWVKCCCTLMVDLGIGSAMLCKTCILAILIPRLDLVISLVGSVSSSALALIIPPLLE 420

Query: 1368 VTFYSEGMSPLTIFKDALISILGFVGFVVGTYEALYELIQPSNAPIFINSTCAFI 1535
VTFYSEGMSPLTIFKDALISILGFVGFVVGTYEALYELIQPSNAPIFINSTCAFI
Sbjct: 421 VTFYSEGMSPLTIFKDALISILGFVGFVVGTYEALYELIQPSNAPIFINSTCAFI 476 (SEQ
ID NO:7)

Hammer search results (Pfam):

Model	Description	Score	E-value	N
PF01490	Transmembrane amino acid transporter protein	223.3	3.5e-63	1
PF01091	PTN/MK heparin-binding protein family	2.0	9.5	1

parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF01091	1/1	192	208 ..	1	17 [.	2.0	9.5
PF01490	1/1	71	451 ..	1	467 [.]	223.3	3.5e-63

FIGURE 2E

Docket No.: CL001062CON

Serial No.: TO BE ASSIGNED

Inventors: WEI, Ming-Hui et al.

Title: ISOLATED HUMAN TRANSPORTER PROTEINS...

```
1 AAAACACGAA AGTCAGATAG TCCCTGTCTC ATCTTCAATC TCTTTATTTG
51 TTATTAGTCT GTTCAGGCTT TCCATTCTCT CCTGATTICA ATCTTGGTAG
101 GTTGATTTCT TCTAGGGATT TCTCCATTTC ATCTGGGTAT TCCAATATGT
151 TGGCAAATAA TTGTTCACAA TAGCCCATTA TGATTCTTTT TATTTCTGAA
201 GCTTCTGTGT TAGTGTCTTC ACTTTCAITT TTTATTTTAT TAGTCTTCTT
251 TTTTCTTTTA GACTAGCAAA GGGTACGTCA ATTTAAATTT TTCCAAAAAA
301 TCAACTCTAG TTTTATTGAT TTTGTCTGTT TTTTCTCTG TTGTCTAITT
351 GATTTATTTT TTTTCTGCTC TTTCTTTTCT TCTTTTAACT TTGGGCTCTG
401 TGAGCTCTTC TTTTCCAGT TTCTGAGGT ATAATGTITAA ACTATTTAAT
451 AGGTCTCTTT CTCTCTTTT AATGTAGGCA TTTATGTCTA TAAACTTCTC
501 TCTTAGAACT ACTTTTGTCT CATTCATAA GCCTTGGTAT GTTATGTCTC
551 CATAGGCAAT TGTCTGAAAA TATTTTTTAA ATTTCCCGTT TGATTTTCATC
601 TTTGACCCAC TGGTTGTTTC GAAGCATAGG ACTGGGTATG GTGGCTCACA
651 CCTATAATCT CAGCACTTTG GGAGGCCGAG GTGGGCAGAT CACCTGAGGT
701 CAGGAGTTTG AGACCGCTG ACCAACATGG TGAACCTCGT CTCTACTAAA
751 AATACAAAAA TAGTCGCGCG TGGTGGCAGC TGCCTGTAAT CCCAGCTACT
801 TGGGAGGCTG AGACAGGAGA ATCGCTTGAA TCCAGGAGGC AGAGGTTGCA
851 GTGAGCCAAG ATTGCACCAC TGCACCTCAG ATGGGGCAAT AAGAGGCCAA
901 CTTTGTCTCA AAGAAAAAAA AAAGCGTGTT GTTTAAATTC CACATATTTG
951 TGAATTTTCC AAGATTCTCT CTGTCAATGA TTTCTAGCTT CATATTATTA
1001 TGGTCTGAAA GAATATTAAT ATGATTTCAA TCTTCTTAAA TTTAAGGCTT
1051 GTTTTTTGGA CTAGCATATG GTCTATTCTA GAGAATGTIT CAAGTGTGTT
1101 AGAGAAAAAA TGTGTATTCT GTTCTGTGTT AATGGAAAGT TCTGTATATA
1151 TCTGTTAGGT TCATTTGGTT TAAAGTGCAA TCCAAGTTCA TTTATTTCTT
1201 ATTTTCTCTC TAGTTGCTCT ATCCATTGTT GAAAGTGGGA TATTGACCTT
1251 CCTACTATTG TGTGTCTATC TATTTCTCCC TTCATGGCCA TTAATATTAG
1301 TTGTATGTAT TTAGGTGCTC CAATTTTGCA TGCATATATA TTTACAGTTG
1351 TGTCTTCTTG ATGAATTGAC CCCTTTATTA TTAACAATAG ATCTTCTCTG
1401 TCTCTTGTGA CAGTTTTTGA CTTGAAGCCT ATTTGTTACA TAATTGTTAA
1451 AGGAAAAGTC TGTAAACAAG AGGTAAAAGG AGAAGCCTAG ATAATACAAT
1501 ACTGAATGAT TGCCATCCAT TTAATAATGT TACTTTAAAA ATTTAGATGT
1551 ATTAAAAAGT AATGTTGCCC TACCCACAG AGTCTTTTCC AGTAGCAACC
1601 ACAGATGATA GTTTGTGTAT GCTTCTGAAA AATTGGAAAG TTTAAAAATA
1651 TGCAATATTT TTATTATAAA AGCAATACAA ACTCATCGAG ATGTGTAAAA
1701 GAAAAATACAG CCAGTGTAAG AATTAGCAAT ATTTACAAAA CCCACAACCT
1751 AAGGGACAGT GCTCTTCGAC TGGACTCTGC CCAATGCCCA AGATCAATGC
1801 CCTGTTCACT TCCTATTCGC AGTCCCGAGC GCCCAGGAAC ATAGTCCTTC
1851 CAGCAGTGGC AGTAATAGGT CGCCAGGTGG TGCTGTGGAG CAGAGCTCCG
1901 GAGCTCAGTG AGAAAAAAGG CGCGGCCGCT CAAGGGAGCA CGTGACCTCG
1951 GCCTCTGGCG TGGGCGGTGG GATCAGCTGA TGAGGTCCGG AAGCGGCTGC
2001 CGGGCAGCAA AGGAGGATGG CGAGGGGCTG ATACTGAACC CGGGAAGGGT
2051 GGGCTGTGCT GAAGCCAGAG CCGGAGCCGG AGCTGGGGCC AGAACCCGAG
2101 CAGTGAGTTC CTCCACTGAC GAGTTCGGGC TGGCGGCGCT CGCCGCTTTG
2151 GGCAGGACCC ACCTCGCCTT CCTCCCGGCG TGGCAGATGC TCCAGGTGAG
2201 GCACGTGGAT CGCCCGGGCT GTGGGTCGCG GACTCCTTGG CGTCCCGGGG
2251 CGCAGCTGCG GGTACGACGC TGACACCCCT CTGTGAATTG GGCAGAGCGT
2301 GGAGATCCCT TGTCCCTCGC GCTATCTCCC TTGACCTCGT GGGGTTGGGA
2351 TCTCACCGTC CTGTTTGACT GACAGGTGGG GGAAACTGGG GTAGATGGTG
2401 AAGATAACCC AAAGGACCAT CTAGGCGCTC TTTACGCTT CGCACAGGTC
2451 TCCCCGTTTC CAGCAATATG CTTGCCCGCT GCGGGAGCGC TGCTTGAGAC
2501 AGGCTCATAA TGGGTCCTTG GGTCAAGAACT GCAAGGACGC TGGGAAGTCG
2551 TCTGGTGCAG CTCCCTCCTA GGACAGTTGG AGAAACTGAG CCCTTACTCC
2601 GGAAGGGGCT AAGGGCTTGC CTAAGGTGAT CCAGTGAGTT AATCGGAGAC
2651 CCGGAGACCT GCGACTAGAA TGAACATGTT CCTAAGCTTC AGCAGCTGTT
2701 TGCTTTTTCG CACACCGCCT CCTGCGGGAA ACTTCACCTG TGAAGGCA
2751 CTCTTTCTG TCCCTTTCTC TTTTAGTCTT CTCCCTTTT AGCTGTCTGC
2801 ATTTTCCACC GCTGGGGTTG GATTGTCTGT GGGTGTGGTT CCCTGTTTGT
2851 TCATTATTTT TCTGCAAACT CATCCTTCTG TAGGTTTGGT TTCTAACCTT
2901 CTGCAATTCT ATGTAAGTCA CACCAAAATA TGAATATGTA ATCGGAATGT
2951 GCTTCTGGGA AGATAGGTGG CTGAGCCGAG GTTGTGGAGA GCCCTGACGT
3001 TAACCTGAAG AATGTAAAGA CCTTGTGCTT ATTTTCTCTG TAACCTGTCA
3051 GATTTGGGAT TGCTTATTTG GATGGACGTT TTGCAGTAT TGAATTTTG
3101 CTGAAGATAG CATCATGGTG CAATGGACAT AACAGAGATT GGGGAATCAG
3151 GATATTTTGT CCTAGTCTG CCGCTTACCT GGCAACCTTA AGTGACTCGC
3201 GTTTGGGTTT CTCAGTCTAG ACAGTGATGG AATTGAATTC TTAAGGGCCC
3251 CTTCTGCTGT GATCTGGATG TTGTGCATCT TTCTAGGCTT GTTTTCTTGT
3301 TTGTTTGTGT TTAATATAGG ATGAGGTCTC ACTATGCTGC CCAGGCTGAT
3351 CTCAAACTCC TGGGCTCAAG TGATCTCTCC ACCTTGGCCT CCCAAAGTAT
```

FIGURE 3A


```

3401 TGC GATTACA GGGGTGTGAG CCA GTGCCCC TGACCAGGGT CTGTTTGTT
3451 TTTTATTCG AGAGATTTTA CCCGCTGTGT ACACTGAGTA TCAGCCTTGC
3501 AACAA GACTT AATCTAATTG TGTAGGAAGC AGTTTCCTCT GCTTATTCCT
3551 CTGTTGCTAT AAAATCCTCC TCCTCTTTCT TCCTATCTCT GTATTATGTC
3601 TAAGCTAAAT ACTAACAGCT GAAAATGATT TTAACTGTCT TGTATTATT
3651 TTAACATGA TCAGGGCCCT CTCTGACTC TTGTCTAGAG CCTCGTTTAC
3701 AGTAATTAAC TTACTTGATC ATTTAGAACA TGTTTTCGTT AAAAATGTTT
3751 TTGAAGTCAA GTGGAAAGGG ACAAACCTTG TGTTTGTCTG GGGATGGAGG
3801 CTACGTGCAG CAGAGGCTGT GCTGATAACA GCTTTAAGCC TCTCCCTTTG
3851 TTTTCTGATT GTATCGTTTA TATTGCGCTG CTCTGGAGTC CTGTTTTTCC
3901 CTTTAAGGCA TGTAGCTGTT CATTGAGCCT CCTCTGTGGG CAGTTCGAAG
3951 TGCTCAGAGT TTTCTCTGTT CCAAAGGCGG TGTAAGAAAA AGGATTCCAT
4001 TTATTTTATA ATATTCTGGA GACTACAACT GGGACCAATA GATGAAAGTG
4051 CCATAAGGGA AGGGTGTTTG GTAAGTGTTA AAAAATTTCA AAGAATTAGT
4101 GGTCTCCTGT ATTAGAACAC TGAGCACGGA GCTCCCTGTC ACTGGAGGTA
4151 ATCGTCTCTG GGCAATTACT TCGGGGATGT TGATAGTGAA ATTAGATTGT
4201 TGGGTAGAGG GTTGGACTAG ACAAAATTTT AAGTTTCTT TCAACTCAAG
4251 AGTCTGTGAC ATTCTAGGAC TGGACTTACT AGCATGTAGA GTGGATGGAG
4301 CAGATGTCCA CTACTAGCA TGTGGGATGG ATGGAGCAGA TGTCCACTTA
4351 TTACTAGCAT GTGGATGGA TGGAGCAGGT GTCCACTTAG TGGCATGTAG
4401 AGTGGATGGA GCAGATGTCT GCTTACTAGC ATGTGGGATG AATGGAGCCG
4451 ATGTCCACTT ACTGGCATGT AGAGTGGATG TTAAGTGGTA GTAAAGTGGG
4501 GCATGTGGGA TGGATGGAGC CGATGTCCAC TAGCACGTAG AGTGACGGA
4551 TGGAGCAGAT GTCCGCTTAC GATGGAGCTG ATGTCCACTT ACTAGCATGT
4601 ACTTACTAGC ATGTGGGATG GATGGAGCTG GCATGTGGGA TGGACGGAGC
4651 GGGATGGATG GAGCTGATGT CCACTTACTA GCATGTGGGA GTCCACTTAC
4701 CGGTGCTCAC TTACTAGCAT GTGGGATGGA CGGAGCCGGT GTCCACTTAC
4751 TAGCATGTGG GATGGATGGG GCAGGTGTCC ACTTACTAGC ATGTAGAGTG
4801 GATGGAGCAG ATGTCCACTT ACTAGCATGT AGAGTGGATG GAGCAGATGT
4851 CCACTTACTA GCATGTGGGA TGGACGGAAC TGGTGCTCAC TTAAGTGGG
4901 TGTGGGATGG ATGGAGCCGA TGTCCACTTA CTAGCATGTG GGATGGAGCT
4951 AGCAGATGTC CACTTACTGG CATGTAGAGT GGATGGAGCT GATGTCCAGT
5001 TTTGTGATTA CTTTGTTTCT ATTTATAACC TTGTCTCAGG TAACTATTCT
5051 CATATTAAGT ACTCCTGCTT TTTCTTTCTT TTTATCACCA CCACTTCCCC
5101 TCCAGTGAGT ATCTCAGTTC TTTAAATGCT TGATATCGCT TCAAAGGTCA
5151 GATGAGTGAA TAGTCTTCTG TTTCTGCTTT TCCTGGCCTG GTGCTGATAA
5201 CCGCTTCCAA AGTGCATGAC TGATTAGCAT TACTCACACC TAGGCCAGCT
5251 TTTCTCTTTT TCCCATAGAG GAACTCACAT GGAATCCGTT TATTTCCATC
5301 CAGGCCCTCT CTTGTTTCCC ATAGAAGAAG CTCTCACTGA GTCTGTTTAC
5351 TTCCACCCAG GCCTTGAAG AATCCTGTAC CTCTCTCCTT TGGCCAGGCC
5401 TTACTGTGAT GAGCACATAA AGGTAGCTC TACTTAATGG GCATGGGGGC
5451 CGATGGATGG GGCATTGTAA ATAGGCTGAA ATAGGAACCA CACGGTGCTG
5501 CATTTGGGGT TGTCTCTTCT TTTATCCCC AAAATATTTC TCTTGAAGC
5551 CTTGACACCC AGGGCCAGTT CTTTTTTTAC TTATTATTTA TTTTATTTTA
5601 TTATTATTAT TTTTTTGAG ATGAGGTCTC ACTATTTTGT CCACGCTGGT
5651 CTTGAACTCC TGAGCTCAAG CAGTCTCTCC ACCTTGGCCT CCCAAAGTGC
5701 CGGGATTACA GATGTGAGCC ACCATGCTTG GCCCAGGGC TAGTCCTTAA
5751 TGACCATCTC TAGTAGGAAA AGCCACTCTG TGCCTTCTT TTCCATGAAG
5801 TTAGGAAATC TGTCTGTGGG TTAAGTATGAT GTTCATGTCA CTTAGATCAC
5851 CATCTCCAAG GTAGGGACCT GGCTTCACAA TCCAGAGTTT TAAATGGAGC
5901 CCATACTGCA GACTTTGTTT AGTGAACCTT CTCACCTTCT GTCTTGAAC
5951 TTCTCTGTAG TAATAATCAC AATTGCTATC ATTAATGAGG GTTTATTATG
6001 TTCCAGGCCAA CATATCTAAC ATTTATTTAT TTTTCTCCTC TATCTTCATA
6051 ACAATATTGT GATGTAGATG TTATTAATGA CATCTTTCAG ATGAGGAGAC
6101 TGTGGCAAGG GGAGATGAAT TAACTTGCTC AGAGTCACAC TACCAGACTG
6151 CAAACTCAGG TGCTTTTTAT TGTGCAGAA ATCCGCTGCA GACCTAATCC
6201 TGCCCAGGCC TCTGGGGCCA GCTTTGTTCT CAGGGAGATT TTAAGGAGGG
6251 TATATAAATT AAGGTGTGGT AGAAAGAATA CTGGACTGGG ATGCTGGTTT
6301 GCTGGACTGT CATCTCAAAT CTTTGATTGG ACTCATCCTG GGGCCTGGAT
6351 GAGTCAGCCT TTGTGTGTGG GCCCTGGTTT TCTGACCCTC AAGAAGGAAG
6401 CTGGAGCTTG ACCTTCTCTA AAGCTATACC TGGCCCTAAC ATTTAGTGAT
6451 CTTATGTTT GGGAGTAAAA GTGTGCGTGT TTGCCTGTTC AGCAGCTGCT
6501 TTGTGCGAAA CCGTCTGAGG TCAGCAGCTG CCCTGTAGCT GTTCTAGCAT
6551 CAGACTCCTA CAGGAAAAAG TCTCAATTTA TGGAAATGTT TGCTCTGGTA
6601 AGTTGGATGG AATTCTATCT GATGCTGTTT TAAAAACAAA TTATGTAGAA
6651 GCCAAACCAT TTTACTTCCC TCACTGTAGA CCACACATAG CAACACAGTC
6701 TGTGTCTTTG TTCTATGTTT TAGAATTTCA TCGACAGAGA GGAGAAAAATA
6751 CATCTGGGGA ATTTGCCGCT GCTCTGAGTT CCAAGTCCA AACCAATGTA

```

FIGURE 3B

6801 ATTGTTTCAG AATAACGGAT GACACTTTTA GCTTGCAAA C AAGGGGCGCC
 6851 AATGCGGTAA TTCTGGTAGG AGGTGAGGCC TAGGGTGTAC CTATCATAA
 6901 AAGATCATAT ATTTTGTGTA GTGCTTTATA TAAATCTACC TATAATCAAG
 6951 ATTACCTAGG AAGCTAGTTA AAAATAAAAC GCCTCTTGCC TGTAATCCCA
 7001 TCACTTTGAG AGGCTGAGAC AGGTGGATCC CTCTACTAAA AACACAAAAA
 7051 ACCAGCTGG CCAACACGGA GAAACTCCAT CTCTACTAAA CTCGGGAGGC
 7101 ATTATCTGGG CATGGTGATG GACGCCTGTA ATCCAGCCCA CTGCGGAGGC
 7151 TGAGGCAGGA GAATCGCTTG AACCCAGGAG GCGGAGGTGG CAGTGAGCCA
 7201 AGATCACACC ATTGCACTCC AGCCTGGGCA ACAGAGGGAG ACACCATCTC
 7251 AAAAAAAAAA AAGAAGACA AAAAGACAAA AACACACATA AAAAAACATA
 7301 GGCTGGGCAT GGTGACTCAT GCCTGTAATC CCAGCACTTT GGGGAGCCAA
 7351 GGTGGGTGGA CCACCTGAGG TCAAGAGTTT GACACCAACC TGCCCAACAT
 7401 GATGAAACCC CGTCTCTACT AAAAAATAGAA AAAAATTAGC CAGTTGTGGT
 7451 GGCGCATGTC CGTAATCCCA GCTACTCGGG AGGCTAAGAG AGGAGAATTG
 7501 CTTGAACCTG GGAGGCGGAG GTTGACGCGA GCCAAGATCG CACCCTGCA
 7551 CTCAGCCTG GCAACAAGA ATGAACTCC ATCTCCAGTA AATAAATTAA
 7601 AATAAATAAA TAAAAATAAA TAAAAATAAA TGCTAAGGTG GAATCAAGTT
 7651 GGGCCAGAAA ATCTATTTTT TTTTCTCTG ACGTATGTTT CATTTAACCC
 7701 AATATATCCC AGATATTATC ATTGCAATAT ATAATCAGTA TAAAGATTAT
 7751 TAATTATGGA GATATTTTAC AATTTTTTTT TTACCAAGTT ATTGAAATCT
 7801 AGTGTGCACA TTTCAATTTT ACCCAAGTGT ATTTCAAGTG TAAGATAGCT
 7851 ATTTATGGCT AGTGGTACT GTACTGGATG GTACAACTTC AGAATATGTT
 7901 ACCATCTATT GATCTTAATC CTCCTTTATT TTGAACAAAC CCAGTCACTA
 7951 AAAAAATGAA ATTGGAATCC TGAAACTTTA GAAGTGAAAG TGTACTTTAG
 8001 AATCATCTAA TGCAAGTTTC TCAATTTCTAT ATCAAAATAA GAAAAACAAT
 8051 TTGGGATTAG AATGACAGCC AGATTATGTT CTCCTGAGTC CTGAATCCCA
 8101 TGCTGTTAAA ATGGGAACAT TAGCATTTGA ATTTATTAGA AAAAATTCTG
 8151 GCCTTGCCCT AAAAAAATAA AATCACTGTA GAATCCCCCT TAAAAATGCC
 8201 CACTTCTGAA AAATTTAACA CCTACAAATT TTTATTTTTA AAAATAGAAT
 8251 AAAATTTATT TTATTTTTAA AAATAAAAAA TCAGTTTGCA CATACATTTT
 8301 CCATATTGCA TCCGTTGCAC AAAGTGATTC CACCTGCTCA TTTTAGTGC
 8351 CCATCTAAAA ATGGCATATT TTGTAGATTG AAGAGCAACA CTGTCTATT
 8401 TATACAGTCA AAACAATAGT TACATAAGGA AAAAAAGGA ATGTTTTAAG
 8451 GTTGACACAC TTAATTTTTT TTTTTTTTTT TTTTTTTTTT GCCATCAAAC
 8501 TTTGAGACTT TTTTACTCA GTTGCTCACT CTCTGAGTC TAAATATCTA
 8551 ATGGAGATTG GGACTTTGTG TTCTGTTTAT TGTCTCAGT AATCTGAAG
 8601 ACAAGCTTGC CTTCAACTCT CACATAGTAC AACCTCATT TAGACAGTTA
 8651 ACAGGTACTA TTAATAATCT CCATAGGGCG GGAAGTGGCA ATTGCAGCAA
 8701 TACTAGTTGC TATCAGATTG CATCAAAGGG AGCCTAAGGG CAGTGTGGCC
 8751 ATGGATGCCA GCATCATGGG GGACAGACAG AGAGCAGGAG GAGGAGGCCCT
 8801 TGGTTTCCAA AAAAGGCCAT AGAAAGAACT CCGGGAGTG GCTCTGCCCA
 8851 CTGTCTGATG CTTGAATCCT TACATAACTG CTCTGAGAAA GGGCTTTTGC
 8901 TTGGATTTTT TCAGGGATAA GGGAACAGGC TTTCTCCAG AGTGATCTGT
 8951 TCTATTTGGA ACAGATCTGT CTTTGATAGA AAGTTCTTCC TTACACCTAG
 9001 CAAAAAATCA GCCCTCTTGA CTCTCCACGT ACTGATCTTA GCCCTGCCTG
 9051 ACCTTTGAGG CCCCATAATA CAAGTCTAAT CCATGTGACA GCTTTTTTTT
 9101 TTTTTTTTTT TTGAGACGGA GTCTCGCTCT GTCGCCAGG CTGGAGTGCA
 9151 GTGGCGCGAT CTCGGCTCAC TGCAAGCTCC GCCTCCCGGG TTCACACCAT
 9201 TCTCCTGCCT CAGCCTCCCG AGTAGCTGGG ACTACAGGCG CCGCTACCA
 9251 CGCCCGGCTA ATTTTTTGTA TTTTAGTAG AGACGGCTTT TAAAGACAGT
 9301 TTTTGTACCC CTCAGATTGC TAGGTGGAAC CTCTCAGTG CTTTCAACCA
 9351 TTCTCATTT AGTTGGTTTC CTACCCCTCT TGATCCTAGT TCTGACCCCT
 9401 GGATATACCA CAATTTGTCA TTATCCCTTT TATAGCATGC TGCTGGGAAG
 9451 AGAACACATT ATCTGGCAAT TCTGAGTTGT GTAACATGTA CCCATGTGTA
 9501 ACATCTAGCA TGTAGAGTGG ATGGAGCAGA TGTCACCTTA CTGGCTGTGA
 9551 GAGTGGATGG AGCCGAGGTC CACTTACTAG CATGTGGGAT GGATGGAGCC
 9601 GATGTCTACC CATGATCTC TCAATGCTTA ATGCAACCTA GAATTGTGTT
 9651 GGTTTATTGG GCATCTTTGA TTATATTATT GCCTTCTGTT GAGCTTATCA
 9701 TCAACAGAAA CTCCTAACGA GAATTTTTTT TTTCTGTTTT CAATATGCAT
 9751 GCAGTTGCTT AGCCACATCT TCCATATCCC GCAGGCTTAT TGGACCTTAA
 9801 ATCTATAGAT TTCAGTTTCT TGTGTAGAAG TTTTCTTAGC AATATTCTGC
 9851 TGTGGGTAA TTTAGTTTT ATACACAACA TAAAGAAAGT TCTCTGCAAT
 9901 GTTTGAGATA AATTGAACAT CTGTACCAAG TAGACAACAG AGAGGTTTCT
 9951 CGTTGTCTAG GGAAGGATTG GGCAATTAAT AAGTCCCTGT ATTCCATCCT
 10001 TTTACCTTCA GTAATATATA GGTGTCAACC TAAAGGAAGA AGTTGAGACA
 10051 CAAAATGCAA TTTTAAACAG TTTACTTGAA CTGTTTACTT GAACCAAGTG
 10101 AGGACAGCTG CCGGGGACAC ACTTCCAAGT TGCCTGGGGG AGTGCCTCT
 10151 TCGGCCTTTG TTACACACAGA TTCTTAAAGG CAAAAGCGAA CAAGGAGAGG

FIGURE 3C

10201 ACTGATACAA AGTGACTTGA CAGGAATTCT CATCAGTTTA CAGAAATAGC
 10251 ATGGATTATT GATGGGCTGT ACATTGTTGG ACTATAGGGT ATGAGTTATG
 10301 ATGTCCAGTG TTAGCATTTT ATGACTTAGT GGTGTCAAGT AGTCTAGAAC
 10351 CCACATAGCA AGTGGCTTCA AGAGGTAATT ATTTAACTCA AGGGGGAGT
 10401 GACACATGAC TGCCTCACA TTTTAGTGCC TCTCTGGACC CGTAATTTAA
 10451 AGGGATTCTT CAGATAAAAA GTTTCTTTTC TTTCTACAA GATTCACTTG
 10501 GAAGGTTCTA TCCTCAGATG CTTTTGTTT GTTGAAGGG ACTAGAAATT
 10551 GGCAGCTTTT TCTTTTTTTC AGTAGAGGCA GGGTCTCACT ATGGTGCCCA
 10601 GGTGTGTTCT GAACTCCTGG GCCCAAGTGA TCCTCCCACTC TCAGCTCTCC
 10651 CAAATGCTGG GATTACAGGT GTGAGACACT GCACTCAGCT GCTGTTTGA
 10701 TAAATAATTA TGTTCAATTG CACCTAGAAT ATTAGTGCTA GAGGGAGTTG
 10751 AGAGATATTT TAGTTTATGC CCCATGCTTT TTGCACATTT AAAAAAGTTT
 10801 CACAGGTACT AAGCAAACTG TTGACAGAGG TAGGCTTGGC GCCTGGGCTC
 10851 CCTGACATAC CTGTAACCTG ATTTACGAGC TTATACCTGT ATAGCAAGAG
 10901 GTTACAATGC TGGTATTAAG ATACTTCAGA GATTTTTTTT TTCTCCCGGC
 10951 CCTCTAGTGA GTTTAAATTGC CCCAGAGCTG GTTGGGCTCC TTGAATTCTC
 11001 CTAGCTCATG AGTAAATGAA GCTCTCATAG ATTTTTAGCC AAGTGGCTCT
 11051 GGCAATGAAG CTAGGCAGGA TCGTCTCTGG GATTTCCAGG TCCTTTGCTG
 11101 GCATTTTGCC AGGTACTTCC CTGTGTGAGT AGCTTGGGG TCCTTCTAC
 11151 ATTGCAATTG TTGAGAGAAA ATGCGATCTC CCGTGGATCT CTCTGGTGCC
 11201 AGACTGGGGT GTTTCCAAAG GAGTACCCTG GCACTGGACC TAAGGAGAGC
 11251 CTTCGGCGGA GCACCATCCT CTGGCAGGTG GTGCTGGGTG TCGGGGCAGG
 11301 GTGGGGTGCT GTGGCAGCAG TTGGAGGTCC TGTCTCTCT CAAGGTAGCT
 11351 GAGATAGAGT GCCCAGGCTT AAGGTGGGCA TCCAGCCACA TGCCGGAGGA
 11401 CAGTCTGACG GGCAAGTAGC TGTGCCAGTC TGCCAAGTGT CGGGAGGATT
 11451 TTTGTCAATT TTTATATTA TGTACCTTTT TTTTGTCACT TGGTTCTTGA
 11501 AGAGCAGGAA GTTGACTCTT TCACTGTGCT GTAATACTCT CTATAGCAA
 11551 CTGGGACTCT GTAGAGTGGT TGTTTTCAGA TTCTGACAGG GGTGAGGAGA
 11601 TAACGTTTTT TGATTCTTC ATCTTTAAAG TAAGGGGCTT AGAGTGACCT
 11651 TCTTTTAGCA TTGATTCTTC ATCTTTAAAG TAAGGGGCTT AGAGTGACCT
 11701 GTGAAGGCTC ATTTAGCAAT GGTCTCTAGG ATCTCTAGGG CTACATCCG
 11751 TCCCAAAATG TGTCTTATTT GTGTATCTTT AAGAAGCCCT TGCTCTCTC
 11801 TTTTGTGTAG TATTAATAGT ATTCTGAGT AAATCCACCC AGGGGACACC
 11851 ACTCTCACCA CCCTCCCAAC ACATTGAAAG GACATTTTTT TCTCTACCA
 11901 TTTTAAAAAT GAGCACATCT ATAAAAATAA AAAGGAAGAA GAGTTGTGGA
 11951 TGGGAGATGT TGAGACGAGG GCCAGGGTGA GCACCTTTCA GTTTCTGGT
 12001 CCTCTCTGA GCTGCTTTCA GCTACCATTT CTCAGTACT AGGTGGGAG
 12051 CAAGAAAGGA GTCCAGGGT CAGGGTATAA GAGTTACTGG TGGCCTCCAG
 12101 AGAGTATAGG ATCAGCCTGT GGTACACAGA GAGAGAAAGA GAACGGCATG
 12151 TGTGGCTCTG GGATTTGGTG GGAGTTTCTG CAGAAATTGA TGATCCAGAG
 12201 GGGATTTCTG TTTTCTTTT TTTTGTGTT TGTGTTGTT TTTTGTGAG
 12251 ATGGAGTTTC GCTCTTGTG CCCAGGCTGG AGTGCAATGG CACGATCTCG
 12301 GCTCACCACA ACCTCCGCT CCCGTTTCA AGCGATTCTC CTGCTCAGC
 12351 CTCCCGAGTA GCTGGGATTA CAGGCATGCA CCACCACGCC CGGCTAATT
 12401 TGATTTTTTA GTAGAGACGG GGTCTCTCCA TGTACTCTAG GCTGGTCTTG
 12451 AACTCCGGAC CTCAGGTGAT CCGCCCGCT TGGCCTCCCA AAGTGCTGGG
 12501 GTTGCAAGCG TGAGCCACCA CACCCAGCCC AGAGAGGATT TCTTGAGTGA
 12551 AATGTGTTCT CTATTGAAG CAGAGGAAAA AGAGTATACG ATGAGAAATA
 12601 CCCAGATTTT CATCCCCCA AGAGCTTGT CATATATAGA TATACGTGTG
 12651 TATATTGTAT ACGTGTATAA TATAGTAACA TACACCTGT ATATACGTAT
 12701 ATATATATAT ATATATATGG GATGATTTAT ATATATATAT ATACAGCAG
 12751 ACGATTGAAC TATTGCACAA GGTCCAAGAC ATTATCTCAG AAAGGAGTAG
 12801 ATAACTCTGA CTAAGGAAT AGGGAATGCG GAATTCAGG AAGCACTTCT
 12851 CTTTCATTTT CCCCACCTCC TCCAAGCAG TGCCTCACT CTGCTTGTG
 12901 TAGCTGTACT CCGGAAAAAT AAGAAATTTA TGAGTGTAGC ACCACGTATA
 12951 CCAATGGGAA GGATGGGAGT CAGAAGTCAA GTGAATCAG CCCGCTCTG
 13001 TGTAATTTTC ACTTTTCCAT TTCCCTTGGT ACCAGGCAT TCTACTCTA
 13051 ATCCATAGTG GAGCTGTGAC AGTGAGCAAC TCTGACAATG ACAGCTTCTA
 13101 CCCAGAGGCG CACCCCAAC ATGGAGCTAA AGGCTCCAGC TGCAGGAGT
 13151 CTTAATGCTG GCCCTGTCCC CCCAGCTGCC ATGTCCACGC AGAGACTCTG
 13201 GAATGAAGAC TACCACGACT ACAGCTCCAC GGACGTGAGC CCTGAGGAGA
 13251 GCCCGTCGGA AGGCCCTCAAC AACCTCTCTT CCCCAGGCTC CTACAGCGC
 13301 TTTGGTCAAA GCAATAGCAC AACGTGAGTA GCTGTACTT TCTCTCTCC
 13351 TGGGTGGGAT TCGTGTCTCT AAGCCTCCCT TGGACTTATT TTTCCCCCA
 13401 ATTTCACTCAG TCCTCACTT TACAGATGAA GGTGAGCAGT GAAGAGAGTT
 13451 GGCAGGTGAC TGCCTGAGA TTTGCCCTTC TGGGCTGCCA CTCTCTAGGC
 13501 AGTTTCTTAC TCCTTTTTTC TTTCACTGT GTTGGCCCCC CAAGGCTGGT
 13551 GCCAAGTGAG AGCTTGACT TAAAAAAGC TTCTACAGAG GACATTTCTT

FIGURE 3D

```

13601 TAATTTAAAA GTGTGTCATC TGTGCTAGAA CCCCAAATAA TTTCCAAGCA
13651 TAATCGGAAGT CTTCCTTTGC AAAGTCTCCC CCCGAATTCT GCCCCATCAC
13701 CAAATCAGTA TTCATTTTGC TGAAGAAGTG GGAAGAGAGA AGAATTTAAGT
13751 TCTGCACITTA AAAAATTTCAG GGTGTGGTAGG AAAGGAAAGA TAGACTTTTGC
13801 ATTCTCCAAA GAGGGCTTAA TCTCTTGTCT CCAGAAACTG GGACCCACGA
13851 CTCAATTTGGG CTGAGTTTGG CCCGCTTCAG GTCTCACTTT CCCCAAATGT
13901 AAAGAAAAAT TGAGGACTCC ACCACAAAGC TATGCTGGCT GTGTGGGGCT
13951 CACCACTTGA ATTAGAAAAT TCAGAGGAAG TTTTGCTACT CCATTGAGTT
14001 AGTTTCCCAG CTACTCCTGA TTTTCAGAGA CCTCTGACTT TTCTCTGTGT
14051 CCCAGCATCT CAGCTTTTGC AGTCTGTGTT ATTCTCAAG CTTAGCTATT
14101 ACCTTTTCTG TGTTTTCTTG TGGACTGAGT GTGACTTAAT GAGAGATCCT
14151 TCATGTCCTA GACTTATGCC ATTCCTGATG ACTGCCCAAG CGGACCATGG
14201 AAGCTTCTGG GCTCATCACT GGAGAAGCTC CCTCTGCCTG CACTGTCTGC
14251 TGGTACAGGG CATTTTCTCT TGCGAAGCTG GGTGGAACTA GAAGAATGTC
14301 TTGCCACATT CCGGCCCGT CACCACCACT AGCTGATTTC TATGCCTCAG
14351 GCTGGAAGTA CTCAACCAGT CCTCTAAGAT TCTGTTTCTG TAGCTTATTT
14401 CTCAGGGGTA TGCTTTTGTA GATTCCCAT TAGCCTGCAG TGGGAGTTAG
14451 CTGGTGGTAG ATTGCTTAGA GCACAGCTGG CAGCAGTGAG GATCACCCTG
14501 CCCCTCTTTC CTCCAACCTT ATCAGCATTG GCAGCCCCCA TGCAGAAGCA
14551 TCTCCACACA CAGCCAATGG CATGTGATGG CTTCCCTTCA GAGGTCATGC
14601 TTGTTATCGT AAGTACTTTC TAAGCTTCTCT TCTCTGAGT TTCTTTTGCA
14651 GTTTTGTCTC CTTTTTGATC TCAGATATCA ACTTGTCTAA GCAATATTTA
14701 GCAGATGAGG TCTGGATTTT TATGTTTATA GAGACATCTC TGAAGTCAA
14751 AACCTACCAA CTAGCAACTT TAGGATAGTA GCTCATAGGT TTTGGACAAA
14801 ATTAGTCTCT TGTTCCTTGG AAATCGAACA AATCAGAAGA TACCTTCTCT
14851 AGGCTTGAT TGTGACATTT TCCAGGGTAT ACTTTGTTCC GAGTTTCCCT
14901 TCCTGCCTTG ATGTTGTGAT ACAGTGTAGG TGACCAGGGA AGCCTATCTG
14951 TAGTTGATGG CAGGTATTAC AGTCCCATCA CAGGTGGTAC AAGATAAAGT
15001 AATTTGCTGG GGCTTAGAGG ACTGGTTGAG TACTTCCAGC CTGGGGCATA
15051 GGATCCACGC AAGGATTTAT ATAGAAAACA TGCCAGGTAT GATTAAAGTA
15101 GAGGTTGATT TGGAGGACCT TCTTAACTTA AATTAATATT TTAATATGTC
15151 GGAAGTGTTA GAGACAAGTT TTTGAGCTGG GTTCTTTTAA TATTTCTGGT
15201 TTGCCCCACC CTTTATCTA GTTTGCGCAA GGAACAAAAT ACATTGGAAGT
15251 ACTTCTACAC CTACTGCACA TATGCATGCA CACACCTGGC TCTTCTAGCA
15301 AGTCAAGGGC TCAGCAAAAA CCCCTAGTTA GGGGGTGCAA ATAGGAACCC
15351 CAAACACTTC CATGAGTTTC ATGGGTTACT TCTTTTATTT TTTTGGAGAC
15401 AGGGTCTTGC TCTGTTGTCC AGGCTGGAGT GCACCTGGCA AATCATGGCT
15451 CACTGCAACC TCCATCTCCT GGGCTCAAGT GATCCTCCCA CCTTAGTTTC
15501 CTAAGTAGCT GAGACTACAG GCATGCTCCT GGCTACTTTT TGTATTTTTT
15551 TTTTTTTTTT TGTAGAGATA GGGTTTGTCT ATGTTGCCCA CTTAGTCTTA
15601 AACGCCTGGG CTCGAAGTAT CCGCTGCCT CGGCCTCCCA AAGTGCTTGG
15651 ATTATAGCTA TGAGCACCAT GCCTGACCTG TGAATTTATT CTTAGTGTGT
15701 TCAGTAGAGT TATTTACTAA CACTTGATGT TACCAAGCTA TTGACTGCTT
15751 CGAAGACAGC CTCAATTTAT GCTGTTGGGC AGATTTTTCT TCTGTTGCC
15801 CCTCTGAGTT CCATTATATA TATCAAGCCT CCGTGCTTCT TCCCATGCA
15851 AACTGAAACC AGCAGACTGA AACTGGCTCT CTAAGGTGAG GCTGGAGTAG
15901 TCATTTGCAA AATGTGGTCT GCACACTTTG TGGGCTTCCC AAGACCATTT
15951 CAAGAAGTCT ATGAGGCTAA AACTCTCTTC ATAATAATAC TAAGATGTTA
16001 TCTGCTTTT CACTTGTGGA TATTTGCACT TATAATGTAG AAGCAATGGT
16051 GGGTAAAAAT ACACTGTAGA ACGAATCAAG GCAGTGGCAC CAAATTTATAC
16101 TAGTTGCTGT TGTATTTTTC ACTGCCACAC ATGCGCAAGG AAAAAAGCCC
16151 TTTGCACTTA AGAATGTCTT TGATGAAACT GTAGGATTAC TAATATTTAA
16201 AAAATTTGAGA CCCTTCAGTA TAGGTCTTTA ATATTCTGTG TGGCAAAATG
16251 GGAAGTATGC ATGAAGTACT TCTATGAGTA CCAAAAATAG TTTACTGTCT
16301 TAAGGCAAAG ACCTCGAGTG ATTATATGAG TTGTCAACCA AACTTGCTGC
16351 CTTTTTTTTT TTTTCATAGA ACTAGAAAAGA ACAACTAACA AACTGTAGGT
16401 CATTACAGCC CGAGTACTTG TAAGACATTT TCTTGAAAT GAAAGAAATC
16451 AGCCCATCAC CTCAGGAAAA ACAATAGATA ATACATATCT GTTGCCACGA
16501 ATAAAAATCA AGCTTTCAAG CAAAATTAGG AAAAAAACC AACTTGTATC
16551 CAGTACCATT AGCTTGATAG CCCCTCTACT TGAAGACTTT TCTGATGAGA
16601 TTAGTGGTGA TATTAACAAA TATGACTTTT TGATATTATT AATATACAAT
16651 GAAGATGTTA ACATTTGGAA GATCTGTGTA AACTCAACCA AAGTATGATG
16701 TTAGGAATTC TGCAATGGTA AAAGATCCAT TGAAGAGCA AGATCACCAA
16751 TGGATTTTTT TTTTCTTTTT TTTTGGAGAC AGTCTTGCTC GTGCACCCAG
16801 GCTGGAGTGC AGTGGACAAA TCTTGGCTCA CTGCAACCTC TGCTCCCCGG
16851 ATTCAAGCGA TTTCTTGCC TCAAGCTCCC GAGTAGCTGG GATTACAGGT
16901 GCCTGCCACC ACTCCAGCT AATTTTTATA TTTTAGTAG AGACGGGGTT
16951 TCGCCATGTT GGCCAGGATA GTCTCAATCT CTTGACCTCA TGATCTGCC

```

FIGURE 3E

```
17001 GCCTTGGCCT CCCAAAGTGC TGGGATTACA GGCATGAGCC ACTGCACCTG
17051 GCCTGACTTT TTTTTTTTTT TAAATACTAA ATGTATCAGG GACTTCTGGC
17101 CTCTTATGGT GTGGTGTGAC TTTTATGCTG TTCACITTTG ATCTTTCTGT
17151 TACAGGGTTT GGGGCTTCTG TTATTATTAT TATTATTTTT TAATTTCTCT
17201 TGTTCCTCTA CCAGTGTITG TCGCTCATTG TTTGGTTTGT CATCTCTGT
17251 TGCAAGTTTG GGATCTGAGT CTTTTTTTTT TTTTGAGATG GAGTCTCCCT
17301 CTATTGCCTA GGCTGGAGTA CAGTGGCAGC ATCTTAACTC ACTGCAACCT
17351 CTGCCTCCCG GGTTCAGCA ATTCTCTAC CTTAACCTCC TGAGAAGCTG
17401 GGATTACAGG CACATGCCGC TATGCCCTGGC TAATTTCTGT ATTTTTAGTA
17451 GAGACGGGGT TTCGCTTGT TGGCCAGGCT GGTCTCGAAC TCCTGACCTC
17501 AGGTGATCCA CCGCTTCGGC CTCCCAAAGT AGTGGGATTA TAGGCATGAG
17551 CCACTGTGCC TGGCCAGGTC TGAGCCTTTA CAGTGGTCAG TTCAGTGGTT
17601 AGAACCAGAC CCAAATACAC TTGAAAGGA TAGAGTGTCT GAAGAGAGTT
17651 GGAGCACCCC TCTGGTCTAA TCTCTGAGAG AAGGGATTCT CAGAAATGTC
17701 AGAGAGTGGA GACTTACAGC ACAGTGGATA AGAGGGGAG CTCTGGAGTC
17751 AGACTGCCCA AATTTGAATC CTGCCCCAGC CCTTTACTAG GTATGTGACC
17801 TTGAGCAAC TCCTTCATCA TCTATAAGAT AAAATCTTAC AGGGTTGTGT
17851 TGGAAATAAA ATAAGATAAT GCATATAAGC ACTGAGATCC TAATAAAAGT
17901 TAACGTGCAT GGTATCATT TCCTTGGCTG TCCTCCACTT CAGATGGTTC
17951 CAGACCTTGA TCCACCTGTT AAAAGGCAAC ATTTGGCACAG GACTCCTGGG
18001 ACTCCCTCTG GCGGTGAAAA ATGCAGGCAT CGTGGTAAGG GTCTGCATCA
18051 GTGGAGAGGA GTGGTGACAA ATTTTAGGAG GTAGCTTTTT GITTGTGTTA
18101 AAATGTACTT GCTTTAAAA ATTTTAAATA GAGAAGCATT TTAATAAAAT
18151 CAGTTGACAA AAAGCGGAAT TCAGACATTC ATTCACTTAA AGATATTTAT
18201 TGAGAGTGTT CTGTGCGTTA GGCAGTGTTC TAAGCTCTTA GAATACATCA
18251 GTGAATTAAT TATTCTGCCC CTCATGGAGC TTAAGTCTTA GTGGAGAGGA
18301 TGTACTGAGA TGGCTCGAGC AGTTTCTGTC AATAATATGA ACTAATGAGT
18351 TAGTTACAGA TGCTGCCCCA TTTTCTACAG TCTCCCATGC CCTGTTCTTA
18401 AATGGCCAA TCGAAGAATC TTATGTCTTC TTTTGTGAT TTACTCCAG
18451 TGACTGCCT GCCCAAAGCC ATTCTGGTTT CTTTCGGAGT TGAAGAGAGA
18501 CTCAGAGATG TGGGTGGCCC TTAGCTAAGT GCAGTCTTTC TTGATCTGGC
18551 ATTGCTGTAA AGATAACTTA CCCGTCTCAC CTCACATCCC TTAGCCAGC
18601 TCTTCCACA GTCAAGGAG CCTTCTATT TCCTGATGTG CACCACTCTT
18651 GGAAACNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNTGTTCT GTGCGTTAGG
19851 CACTGTTCTA AGCTCTTAGA ATACATCAGT GAATTAATAA TTCCTGCCCT
19901 CATGGAGCTT ACTTCATGGT GGAGAGGATG TACTGAGATG GCTCGAGCAG
19951 TTTCTGTCAA TAATATGAAT TAATGAGTTA GTTACAGTAT GTCTGCCCAT
20001 TTTCTACAGT CTCCCATGCC CTGTTCTCAA ATGGCCAAC TCAAGAACTT
20051 TATGTCTTCT TTTTGTGATT TACCTCCAGT TGACTGCCTG CCCAAAGCCA
20101 TTTGCTTCT TTTGCGAGTT GAAGAGAGAC TCAGAGATGT GGGTGGCCTT
20151 TAGCTAAGTG CAGTCTTTCT TGATCTGGCA TTGCTGTAAA GATAACTTAC
20201 CCGTCTCACC TCACATCCCT TAGCCAGGCT CTTCCACAGC TCACAGGAGC
20251 CTTCTATTCT GCTGATGTGC ACCAGTCTTG GAACAGACTT ATCTTATGTC
20301 CTTCTTCTCC ACGTGACTAA ATCTCTCGAA AATGTGCTAA AGTTGACTGA
20351 ACACCATCTC CACAGAGCTA ATCTCCATGA TCAGTGTCTT CTCTCTAATG
```

FIGURE 3F

20401 CTGGGCCCAA AGTTTTGTCA CTGAAAAC TGCTAGTAGC TTTTAAATCC
 20451 TTTGTGAAC GAGTATCCAT TGGGTTCACT CCTAATTCTA CCTACTTTTC
 20501 TCTCTCTCT TCGCTGCAA TATCTGTCCC CAGATGGGTC CCATCAGCCT
 20551 GCTGATCATA GGCATCGTGG CCGTGCACTG CATGGGTATC CTGGTGAAT
 20601 GTGCTCACCA CTCTGCGCG AGGTGAGAGC CCTCTGAGCC ACCTCTCAAG
 20651 TGACAGATTG TCCTTTTGGG TTCTGTATC AACCTGAAA ATGAGCACTG
 20701 ATGCAGACCA CTCTCAATTC TTTACACTGG CTGGAGGTAG CAGCTTATGA
 20751 TTGCAGCGTT TTCCTTTCCC TGTTATTTT TCGGTTCTTT TCTGGCTCAT
 20801 TATCATCTGT TAAATTTACT TATGCCAGT GGGTACTACA TTCTAATTTT
 20851 ATGGGCGTTG TAATATTTAC CCAATTGAAA TGATTCTACC AGATGGTTGT
 20901 TAATTATAAT AAAAGTAACC ATCCTGTGCG CTGAATACTT CTGATCTTTG
 20951 AAAGCACGAG ATACAGGACT CAGAGTGGTA CCTCCAGGTT GAAAGATGGG
 21001 AACTGGCCCA GGTCTCAGTG GCTCTTTTGG TTCTGTCAAT GTCAATGTCT
 21051 AATCCACGTG CTCTGTCTTT CCTCTTCCCT CCTACTCTTC CAGGCTGAAT
 21101 AAATCCTTTG TGGATTATGG TGATACTGTG ATGTATGGAC TAGAATCCAG
 21151 CCGCTGCTCC TGGCTCCGGA ACCACGCACA CTGGGGAAGG TAACTGATTT
 21201 CCTCCTTCCT TTCAACTGTG GCCTCCAGT GTGAGGCCCT CAGATGGGGA
 21251 GGTGCAACGT GGGAGACAGT GTAAAGCGTG GAAAAGAGTG TGTGGGTC
 21301 AGTTGCCCTG GGCTGTGGCT CAGCTCTGCT GGTAGTAAGC TGTGTGACCT
 21351 GGGGCTGGGT AACCCCTTTT TTCTTGGGT TTTAGTTTTT TTATCAGGAA
 21401 AGCATGGGGC CTGGCCTGAA TGGTCTCTAG AGCCATTCCA GCTTTGGCGG
 21451 TCTATGACCA GTGATTGTTT TTGATTCAT CATTTGTTC ACAAATGTAT
 21501 TTAAGCACTA TCTTATAAAT GGAACAAAAC AGTTCTAGGT AAGAAGGGAA
 21551 GATTTCCTGA AGTAAATTAT GTGGTTCCTA CCTCCAGAG GCTTGTAGTC
 21601 TGTGTAAGGA AAAAGAAATG TGGAAGAGAG AGCCGGGGAA CAAGATAAGA
 21651 GACCAAGTAG GGGAGACACC CATAAGAGA AAGTGTCTAG AGCTAGGAGT
 21701 ACACCCCTCAG TGCTCAGAGA GAGAGGAAC TTAAGGATTC TCTTGTCCGG
 21751 TGTGCCAGAT GAGAAACGCA CATGAGAGAT AGGAGCAAAG AAGGCTTCAG
 21801 GAGAAGGTGA GATAAAGTAG AGCAGGGCGT GGAGATGAGT TTGGAGGTGG
 21851 GAAGTATTTG CAAATTTCTC GTTATGGTAA CTCTTCAGTG TTTGGAGGGA
 21901 AATATTATGT TTGTTTTCTA CATTTAAATG TAGGAAATTT ATACTATCAA
 21951 GGGCTAAAAA TTCTTAAAAA AAAAAAATAA GAACCACTAT AAAACTATGT
 22001 TCTCTAGAAA AGTTCTTTTT TGTGTCTATA GAGGAACTT ACTTTCATTC
 22051 ATAGTCACCT TTATCCTGTG ATGCAAGTTA TATAGTTCTT TTGCCAAAT
 22101 TATTTTCTGT AACTGGGAGA AGCTAGATTG CCAGGTGACC ACCATGAGTT
 22151 GGGTGGTTGT TAATTTCTTC TTCCATTCTT TCTTACTACT TCCTTTCTTC
 22201 CGCCCTCTCT CCCTCTTTTC CTCTCTTCTT TTTAATAAAA TGTGTGCTAT
 22251 TTTAATGCGT GCTCATAGTA AAAACTTTGT TTTGATCAAG ATAGGACATA
 22301 AAGTAAAAAG TGAAAGAAAA TTTTGGTCAC AGTTGATGAG GTAGCCTTTT
 22351 GGAATTTGCT GTATAAGTAG AAACATACAC ATGTTCTTAA AGTTTTTTGC
 22401 ACAGATTGAC CATACTATGT ATACTGTTTG GAAACTTGCT TTTTCCCTTT
 22451 AAACGTCTGA GACGTTTTTC TCTATCAGCA CATAGAGATT TAACACATTTC
 22501 TTTTAACTG CTGTGTAATG TTCCATTTAA GAACGGTCTA TAAATTAATC
 22551 ACTCTGCTTT TGATGATCCT TTAGTTGTTT ACCAGCTGCT ATTGTTCAAC
 22601 CAGCAGCTGT TTTTGGTAC ATCAGTTTCT GTGTCCTTAA TGTGGGACTT
 22651 GGTGGTCTCT TATATCCAAG TTATAGAGAC AGTGAAGGGG ACTATTTCTT
 22701 TGTGTTTTAT GTCAAGGGCT CCCTGTAAC AACAAAAAAG TGTGAGATGG
 22751 GATAGGTGGG CAGATGTGTA GAGAGGATGC TAAGGGGCTG GGCAGTGGTC
 22801 ATGGTGTCTG TGCATGTGTC TCACCTCATG CAGCAATCCA GACGAGAAGC
 22851 CAGGAAGGGG ACGTCGGAAG CCACACAGAT AGCACCTCCC TCACCTTCTT
 22901 CCCAATGCCC CAGACCACTG GCACCTAGCA TGGTTTCTTC TCCTGCCAGG
 22951 GCATCTCGTC CTTGTCACTG CCAGGAAGGG TCTGTGATGG CTTGGGAAA
 23001 AGCACTGTGA AAAAAACACT TAATGGGCAC AATGTACACT GTTTGGGTGA
 23051 TGGGTACACT AAACGCCAG GCACCTACCA TATGCAGTAT ATCCATTTAA
 23101 CAAAACAGCA CTTGTACTCC CTAATCTAT TAAAAACAA AAACAAAAAA
 23151 CACCTCCCC TCTGGGAGCA TTGCATTTGT ATTGTAACAG TCTTTGTATT
 23201 CCTTCTCTCC CCACCTCCAG ACGTGTGTG GACTTCTTCC TGATTGTAC
 23251 CCAGCTGGGA TTCTGCTGTG TCTATTTTGT GTTTCTGGCT GACAATTTA
 23301 AACAGGTAGG CACCTGGTTA AAAAAAGAAA AAAAAAAGAA AACCAAGCGG
 23351 AGAATGGCAA AAGATGATTG AAGTTTTTGT TTAGGATTTT TTCCAATACA
 23401 GCTTTTGTCA ACAAAGAGT TAAAGTTTTT ATATTTTACA TAGATCTACG
 23451 TCTTCTATT GATTCCCAGT GAAAGAGCTC GGGCATAGAG AAACCCGAC
 23501 ATGTCTTGTG GACCTCCTG TCCTAGGTAC ATATGATCAA ACCTAGCTCA
 23551 GACAAATGGG TTGCTGATGA TAGTCGTGAA GTTCTCTAAA GATGGCTCAC
 23601 TGGCCACAGA TTCTAAAAGG CCTTGTTCAC ACACCTGAGC CTTTCTCTCA
 23651 GAACCTCTTC CAGCAGAGGA TCCACCGGCC TCTGTTGTTT GAGAGGTGTT
 23701 TCCGTTTTCT TCCTTCCCCT CATTCAGGTT GATAGAAGCG GCCAATGGGA
 23751 CCACCAATAA CTGCCACAAC AATGAGACGG TGATTCTGAC GCCTACCATG

FIGURE 3G

23801 GACTCGCGAC TCTACATGCT CTCCTTCTCG CCCTTCTCTGG TGCTGCTGGT
 23851 TTTTCATCAGG AACCTCCGAG CCCTGTCCAT CTCTCTCCCTG TTGGCCAAACA
 23901 TCACCATGCT GGTGAGCTTG GTCATGATCT ACCAGTTCAT TGTTGAGGTA
 23951 CATGCCATGCT CCCTCTCCTA TCATCTTGGT TCAATATTTT AAAAAAGCCA
 24001 GGCCTGGTAG CTCACGCTCG TAATCCGAGC ACTTTGGGAG GTGGGGCGCG
 24051 GTGGATCACC TGAGGTGAGG AGTTTGAGAC CAGCCTGGCC GTCATGGTGA
 24101 AACCTGTCTC TACTACTAAA AATAGAAAAA TTAGGCAATGG GGGTGTGGGC
 24151 CTAATCTCAG CTATTTGGGA GGCTGAGGCA GGAGAATCGC TTGAACCTGG
 24201 GAGGCCGAAGG TTGCACTGAG CTGAGATCAT GCTACTGCAC TCCAACTCTGG
 24251 GCAACAGAGC AAGACTCTGT CTCAAAAAAA AAAAAAAA AAAAAAAA
 24301 TATATATATA TATATATATA TATATATATA TATATAAAT ATATATTTAC
 24351 ATATATGTGT ATATGTTATT ATATTTTACA TATATATTAC ATATATATATA
 24401 TACATATACA TGTAAATATAT ATATGTGACA TGCATAATAT ATATATATATA
 24451 ATGTATGTAA TATTTATATA TTATGTATAT ATATACATAA TATATATATG
 24501 TAAGTGAAGT GTAAATAGTT ATATGTTACT ACTGGTATGT CTAGATTAGA
 24551 GGTTCGTGTC TTGGGCCCTG TTGACATTTT GGGATGGATA AATTCTTTGC
 24601 TATGGGGCTG TCCTGTGCAT TGTGGGGTGT TTAGCAGCAT CTCTGGTCTC
 24651 TCTCATTAGG TACCAGTAGC GATCCCTCCA TGAGTTATGA CAACCAAAAA
 24701 TGTCTCCAGA CATTGCCAAA CCTTCTCTGG GGGCAAAATC GCCCCCAAC
 24751 CCAGGGGGCA CTGGTTTAGA CTTTTTCAA TTAGATGGTT AATTCATGAT
 24801 CATTGTATAC AGTTGAAAAA TAGAGGAAAA TGTTAAGATT AAAATAAAAA
 24851 ATAATTTTTC TAACCTGTAT TTAGATAAGT AATTCCTTAT CAACTCCAGT
 24901 TAATTTTTAT TTGTCAAAAT TATAAATTCA CTGTCTCTT GCCCTCACTT
 24951 AACCCATGCA GGCAAGTCTG TGGGGTGGCA TGAGAGAGAA CATCTGTATA
 25001 CAGATGGGTA GAAATCAGG CTGAGAAAAA TGTGCCCTTA AACACTATGG
 25051 CTGTTTGTGA AAATGAGAA GCATTTTCTA AGGCTTGAGA AAAGGAAAAA
 25101 AGTAAAGGCG GGTAAATAAA AGCATAAATT AAAAAAAA ATACTTAAAT
 25151 TCAGTTCCTC AAATAATTCA TCAGTACATA TTCATTAATA TGCAAGCAAC
 25201 ACAAAATCCT CTGGAATACC ATGTCCCTCC TAGATCTGT GAATGGATT
 25251 CCGGCTGTAT GTGATTGGTC TGTCTCATTG TGTCTCATTG CTTGCTTTT
 25301 ACAGCCCATG TAAGTATATT GAGAAATACA TTGAAATATA TTTTGTTTT
 25351 ATTTTGAATA CATAATTTT AGTGAGCCTT CCAGGTTAGT ATTCCTGGCT
 25401 TTCCACCTTA AAAATGCCTT GTCACATTGT ATCAGAGCAA GGAGATTTCG
 25451 CTACCTTGTT GTTGTAGTT CTCACCTCAGT GGTATCAGG CCATGGATAA
 25501 TGCCATTTAT TTAACAAGTC TTAAGACAGT GAGATGTTCT TATACATTCC
 25551 TTTTATGAT TATTTACAGT TACTTAAGAT ATTTGCTAG ATGTGTAATT
 25601 TTTTGTGGGA CTTGTATAAA GAATTTTGTG TATGTTACCA AGTGTGTTTT
 25651 GCTGAAGAGT GTGCACTTTT TTTATAGAGA CTCCCTTCCA CAGTATTTAA
 25701 TCCAAAGCTT ATTTCTAAT TTCTATAGAG GATGTATTAG TATTTATTTA
 25751 GTGCCCATTT TTGCATCTT ACTAATACTG TAGTATTAT TATAGTATA
 25801 ATTAGTATTC TTGATATATG TATGATATAT CAGGTGAATA TTATCTCAAT
 25851 ATTATATTAG TATTATTTTG GTCAGTCTGA CAGGTGAATA TATAGTATA
 25901 TTCTATACAG CTGCTTAACA GTGACCCAGT CAGGTGAATA TATAGTATA
 25951 AGGAAGACAG TGGCTCATAA AAAGCAGGAC TTCTGTGCT TATAGTATA
 26001 CATTATCAAT TTAGATTAAC ATTTTGTCT GTGAGTATTG ACTGTTTTT
 26051 CATCACATTA AGTCAGATGA GTGGCAGATA TTGACTCTTC TGCAGAACTA
 26101 TTTTTTTAAA CATAAAATAA ATAGTCTTTC ACGTCCCTT TTTATGAGAC
 26151 TGGAAGGAGG GCCTGTGAAT GCACTATAGA CCAATTTCCAT TACCAAAACA
 26201 GGAAGGGTAA ATATTTACCC ATGATTTCAT GTTGTAATTA GAACCTCAA
 26251 GAGCACCTAG GATATACTCC ACTGTACACT AAAGATGAAT TGAGTGAGG
 26301 CACCCCATTA GACAGACACA CCGTCTGCAT TTACTTTTAA ATCATTGAG
 26351 GTAGTCGGGC TGTGTTGGTT GCCATGATCA GAGGGCTGGG ATAAAGATTT
 26401 GGGTCTTTAT AGCGTCTCAG TCCCAACCAA CTGGTAGTAT CCATCCAGAG
 26451 TGATGTCTAT GCATAGTACA ACCAGGACAC AGAGCAATGT CTGCATAAGG
 26501 GCAGCCCTGC TGATTTCTTG AGAGCAATTC TGAGTCTTCC TCTGGGCTTA
 26551 GCCAGAAGTT GTGCTGTGAT CAAATAGTGC CGTCTGCCCT GAGTACAGCA
 26601 TGGGGGAAGT GGTTTGGCTG TGTTTTGATG TAGTCACTGC CCATAGTGT
 26651 GTAGTGTCTT CATTTTGATG TGTCTATACG CTAAGAGATGC TCCCTTTAGG
 26701 TCAATTTTGT TGCCGCTGCC TCTGCGCTT GTTACTACTG TTCTGTTTTG
 26751 GCATTGTGCC CCACTTACCA TGAGGATTCC CCTACTGTTT AATGTTTCTG
 26801 AATTTTTTCC CTAATCCTAA GCATGTACAT GACTGTTCTT CTTGCCCTC
 26851 ATGCACGTGC CATTGTAGGT AGCAGACCAA GGTCTTCCAC AGAGAGCAGG
 26901 TTTCTCTCTG TCTTCAGCAT GTGGAGTCTC AAATGGAAACA GTTCTGGGCA
 26951 GAGTGCTTTG CACAGAGGGT GCTCCCAATA AATGTTTTAT CACTGCATAT
 27001 CGTGTCTTCT GAGATGTATT TTTTATAGT TATAACAGAT TTCTCAAAA
 27051 AAGAGTACAT CTCACAATCC ATGTGTACCT TTAACAGCAT TTTCTCAAAA
 27101 TACTGTTATT ATAATTGATA ATATGGTAAG ACCTCACTTA ATATCATTGA
 27151 TACATCTCTTA GAAACTGCAA TACATTAAT GTATGTATAG CGAAATCAGT

FIGURE 3H

```

27201 TTTTCTCA TCAATGTTAT AACAAACAG CGTTGAAGGA AGTGACTGTA
27251 CGTCATTTC CTTAAAGTCT CAGTTTCCAA GAACCTATTG ACGACAAGGG
27301 AGGACTTACT GTGTTGTAGA ATTGAGGAGA TATGTTAATA ACGAGCTGAT
27351 TTTAACACTG ATGTTTCTTT ATAAATTAAT CTTTCTCAT TTAGTTGGTT
27401 GGGTCAGTAG CAATCAGTAA GTATGTAGAA TAATACACTT CTTCTGCTGG
27451 CCTCATTCCC ACAATATCCC CACATATGGA TTGTGAAATT CCCAGTCTGA
27501 TACCTGAATC TGATCTGATG TATGAATAAG AGCAGGAGTC ATTCACTAAC
27551 CAACAGATAG CACCTGTTTC CAATAACTTA GGTTCATTTC GTGACTCAGG
27601 AATAATTACA GGCCACTCTT GCTCTCAAGT CCCATTGTAA AGGAAAAATA
27651 CCTATTACCC TGCTTTCATT CCAGGTATTG AAATGCTTCT TACAAAGGGA
27701 TCTAACAGAT TTCTTAGCAG GGGCCACAGG AAACACATTT ATTTAATTTT
27751 TTTATTTTTT CAAAAGCAAT ATTACTGCTT TGAATCTTT TGAAATGAG
27801 GCTGTTATAG AGCTTAATAA TGGATCTCCT TTTACTTGCC TGAAATTATT
27851 CTGAAGCCTG TTAAGAGCAT GCCCCGTATT ATCCAAATAG CCATACAGTT
27901 AAATCAATT TAAACATTG TAAAAGGCTG TTTTAAACATC AATTTTTTAT
27951 TTAATTGAAG CAACATACAC ATGTGGTTTA GAAAACCAAA TTGTAATAAG
28001 ACAGCAGCTT TGAATCCCTC CTCCCCACCC TGCCCTTCC ACACAGTCTG
28051 TTAAGGAGAT CTGTTGTTCC GTGAGGAGT TTGTGACTAT ACCTCTGTCT
28101 TAGTCAGGGT TCTCTAGAGG GACAGAACTA ATAGGATAGA TGTACATATA
28151 TAGGGGAGTT TATTAAAGGA TATTAACCTA CGCAATCACA GGTTTCCACA
28201 ACAGGCCGTC TGCAGGCTGA GGAGCAAGGA AGCCAGTCCA AGTCCCAAAG
28251 CTGAAGAACT CGAGGTCTGA TGTTCAGGG CAGGAAGTAT CCAGCACGGG
28301 AGAAAGATGT AGCCTGGGAG GCTAAGCCAG TCTAGCCTTT TCACATTCTT
28351 CTGCTGCTTT TTTAATCTGG CCACTGGCCG AGCTGGCAGC TGATTAGATT
28401 GTGCCCATCC AGATTAAGGG TGGGTCTGCC TTTCCAGTC CACTGACTCA
28451 AATGTGGCAA CACCTCACA AACACACGCA GGAAACAATAC TTTGCTCTCT
28501 TCAGTGCAT CAAGTTGACA CTCAGTATTA ACCATCACAA CCTCTTCTT
28551 TATACAACT TAACATTGTA CTGCAGTTA ACAGTTGCC TTTTCTGCG
28601 CCATTTTTTA AAGCATTCTT TTTGCTCTC CTCCCCACAT GTTCCAGCAC
28651 TCCTGTTGTG TGCTTCTTGG TAATACTTTG AAAGTGCTCA AGTTCATTGA
28701 TGAGAAATTT AAAAAAGGAG AGAAAAAGAG AGGAAAAAGG AAGAGAACCA
28751 ATATAAAAAA GTACCACTTT CTCTTCCCTT CCAGCTTTAT CTTTGATGTT
28801 TATGTAGTTG TATCAGAGTG AATATATAAA TTAATAATTAA AATTTTTTCT
28851 CACTAATATT TCGTAAGTAG TTTTTCATGT TCTACTTAGT TATCTCAATT
28901 TTTACTTTTT AATAGTGCAT ACTGGCCAGG CGCAGTGGCC AACGCCTGTG
28951 ATCCCAGCAC CTTGGGAGGC CGAGGCAGGC CCCTGTCTCT ACCAAAAATA
29001 TTTACAAAAC GCTTGGCCAA CATGGTGAAG AGCTGTCTCT ACCAAAAATA
29051 TACAAAAAAT AGGCCAGTGT GGTGGCAGCG ACCTGTAGTC CTAGCCACTC
29101 AGGAGGCTGA GGCACGAAGA ATTGCTTGAA CCCAGGAGGC AGAGGGAGGT
29151 TGCACTGAGC CGAGATCTCG CCACTGCACG CCAGCCTGGG TGACAGAAGG
29201 AGACTCTGTC TCTAGATAGA TAGATAGATA GATAGATAGA TAGATAGATA
29251 GATAGATAGA ATATATATTC TGTGGTTTAG CTGGTATGTT GTTAATTACT
29301 TAACTGATCC CTTGTTTGGA AGCACTTATA TGTGTTTTCG TTAATTATT
29351 AAACAGCTTT GCTCAGTGT TTTTCTCTT TGATTTTTTT TCTACTTGAA
29401 TTCAGTTTCT GGGGATGGGA TTACCTAGTG AAAGAATGTG ACTCTTTTTA
29451 TGCAAGCCCC AACATTTGAG TTTTAAATAG TACCTGGGCG TTGTCTTTCC
29501 CCCAAACAAG TGGGTTTTTC TTAGCCTGAA GAGAAAAACA TACAAAGGTT
29551 AAATGTCCCT AAATCATCTG TCAGGTATTA GACTTTCTTC CTTTAGAGAA
29601 TCTTGGATT GTTAAAAGGT ATGACCTCTC CGATTGAGG TTTCAATCTT
29651 GAATTTCTGT ATAGCCTTTT GCTTTGTTTT GCTTTCTGTC TTTGAGGA
29701 TCCCAGACCC CAGCCACCTC CCCTTGGTGG CCCCTTGAA GACCTACCTT
29751 CTCTTCTTTG GCACAGCGAT TTTTTCATT GAAGGCATTG GAATGGTAAG
29801 AGCTGCACTG TGAATTTGGG TAGTGTCTC TGGTGCCCTT GGTGTTCTCC
29851 AGTCTGTGTT CAAGGAATGC TGAGGAAACA TTGTTAGAAA GTATCTTCTG
29901 AGGCCAGGCA TGGTGGCTCA CGCTGTAAT CTCAGCACTT TGGGAGGCT
29951 AGACTGGTGG ATCACTTGAG GTCAGGAGTT CGAAACCAGC CTGGCCAACA
30001 TGGTGAACCC CCATCTCTAC TAAATATACA AAAATCAGCT AGGCATGGTG
30051 GCACAGCCTC ATAATCCAGC CCACTCGAGA GGCTGAGGCA GGAGAAATTC
30101 TTGAAGTGGG GAGACGGAGG TTGCAGTGAG CCAAGATCAC GCCACTGCAC
30151 TCCAGCCTGG GTGACAGAGC GAGACTCTGT CTCAAAAAAA AAAAAAATAA
30201 AAAGAAATTA TCTTCTGTAA CTCAGTGTC AGTTAGTGAA TAGTGTTCG
30251 GGGATTCAT TGAGATTTCC CAGCTTCAAC TTTTCAAGC AAATTATATG
30301 TAAATTTTAA ATGTTTACAT TCAAGGCCCT TTCCTGCAC ACTCATCTCC
30351 TATGTGTGCA GTAAGGAATA GCATATGGCA ATCAGGAAGG CAGGCTCTAG
30401 AGTCAGACTG ACATGGGGGT AAGTCTGGC TCTGCCATG AGTAGCTGTG
30451 TGACCTTGAG CAAGGGCTTC ATCTCTTTGA GCCTTCATTA TTTGTTCTG
30501 AAAAGTGAGC TTAATGATTC GTAGTTATTA GGATTAATG AGATATGTGC
30551 AAAATGCTTT GCACAGACCC TGACACATGG TAAATGTTA ATAGATTTT

```

FIGURE 3I

30601 ATTTTATTAA TAATGTTATT TTATTATTGA ATCAATAAAT GCATGAATAA
 30651 TTTCTCTGCC CTACAACATT GGTTTGGTGT ATTTTCTGCG TGCAAAAGAG
 30701 CAGCCTTAC TCCTGGCTCA GCATTCTGTG ATTTACCAA ATGCTTTTCC
 30751 TAAAAAGGAA TCTACCCCTC ACTTTTACCT AATTTGAAAT TTATTTTGAT
 30801 TTTTCAATAA AATGGTACAA AACTCTTCTC TGAAGAAAGT TAITCTTGGC
 30851 CAAGGGCGCC ACAAAATGGGA AGAGCCTCCT CCGGCGACGG CACTCTTTTC
 30901 TCTTTTGTGA TTTCCGAGAC ACCTTATTGT CTTTAAAGAA AACCTAGAAG
 30951 CTGTACACAT TTTTGTCAA AATCGTGAGA AATCACCCAG GTGGTGTTGA
 31001 TGGAAACAGT TGAAGCTCTG ACATCATGGG GGAGGCTTTG GGAGTGATCA
 31051 CATGTAAATC ACTGGTCTCC CTGAGGTTTT ATACCTTGCC CTGTGCTCTA
 31101 TCTTAGGGCT TTTCAATTGCC CATGAAGAGT GTCTCACTGT AATTGAGAAA
 31151 CACAAATGGT TTCCTGTGTC TGGGGCAGAG GCCTGAGTGG GCCCATACTT
 31201 CAGCAGTGAG AAAGAGATCC CAAGAATCA GGACTGGAAA GAAGAGGCTG
 31251 AGAAAGTGTG GAAAGATGCC CAGAGACCTT AGGTTCTTGG GCATCCTAAG
 31301 GGACCTTGTG CTAATTTTTT AGTAGCTTTT CCTAACAGCA CAGCGCAGAA
 31351 ATGTTTGTCT TGGTTTTATT ACCCAAGACT TGTACACAAA GTTATTCTGC
 31401 AAACATCATT TGTTTCAAG ATTTCTTTGT ATTTCTATTT TTTTACAATA
 31451 GAGAGAGAAC ACTGCTAGAT TGACTCTTAG TTTTGGATCT AGGGCTTGT
 31501 CATTGCATCG GGGTAAAGTG CAGGCTGCA CACTGTATTCT ACCGTGTGCT
 31551 CTGTGTTTCA GCAGCTGTCA CAGGCCAGAT ATGGGCTCCC TGCCCTCTGG
 31601 CTCTTTGTAT TCTTGGTATG ATGGAACTG ACAGATACAT TTAGAAGCTG
 31651 TTTCTAATGG ATGTGGATTA ACTGAAGCTG GAGCAGGTGG TGGGGGACAC
 31701 TGGGGGCCCC TGGGAGACC AGGCTGCAG CAGGATCATG TTGGTGTGTG
 31751 CAGCACAGAC CTATTGTGCC TCATGGCTGC AGGTGCTGCA GGAAAATAAG
 31801 TGTAAATAACA AGATCAAGT TGTGGGACCA GCTAGACATT GCCCAGTGT
 31851 GCTCTTCTCT ATTAGCTGTG CCAGATCCAG TTGATGGCAG ACTGTGAAGA
 31901 TCCTCACAGT CACACCTGCC ATTTTCCCAA TTCCTAAATA AACTCATTTT
 31951 CATAGGGGCC TTTCTTTTGC TTTTACCAA GTTAAAGAA GTCTCTCTCT
 32001 ATGATGAAGG AGGTAGCAAA GTGCTGTGAG TTTGTAATAGG AGTCTGGAAA
 32051 CTGGGCATTG TTAAGCACCC ATGTCTTGAG ATCCTCTACG AAGTCATGCT
 32101 TTTCTTTCGA CTGCACTTTC TCTCACTTGG AGGTTTAACT TGTCACTGCT
 32151 AGTGCTTGCC CCTGTTGAGC TAGAGTGGCA GTTTTCCCTG ACCTATATTT
 32201 TGCATTCTTT AGACAGGTTA GCAGGAAGCT GTGATCTCAG GTTAGATGCC
 32251 AGGTGGGCAT GACATGAGAG GGCCTTCTGG TTGCCATGTG GCCTCACTCA
 32301 CAGGGCAAGG GACATCCCA GCCTGCAGGG ATCCTGCAGC AGGAGGACAG
 32351 AGCACTGGCC TGAGCCAGGA GTCCCTGGGCT CGTGTCTCTG CCACCTTAC
 32401 TTTTAGGACT TCCTAGCTAG GCAGTGGGCT GCAGAGTCCC TTCTAGTCCC
 32451 AAGAGCATAA CGTCTGATGA AATAACTTTA TTTAAAGAGC AGATGTGCTT
 32501 CTGGAGAAAT CTGGGGATAA AAGAGTTACT TTTTCTCGA GGTTTTTTTT
 32551 TTTCTTGGCC ATTAACTTTT CTTTTTCTG CACTTCTCT CCTCTCTCAC
 32601 TACTCTCTCA TAGGTTCTGC CCCTGGAAAA CAAAATGAAG GATCCTCGGA
 32651 AGTTCCCACT CATCCTGTAC CTGGGCATGG TCATCGTCAC CATCCTTAC
 32701 ATCAGCCTGG GGTGTCTGGG GTACCTGCAA TTTGGAGCTA ATATCCAAGG
 32751 CAGCATAAACC CTCAACCTGC CCAACTGCTG GTACGTGGAG GGAGGATGGA
 32801 AACCTAGGAG CACTGGATAT TTTAAAAAAC TAATGGGTCA CAGTGTGGAT
 32851 TCTCCCTCTT ACTTATCTCT TAAACCAGCC CACTTCACTC TAGCCCCACA
 32901 TCCCTGCCA CTGCCAGCCC TCACTGGCTG CCCTGGACTG CATTTCTGTT
 32951 GGGGAATTCA TGTAGAGCCT TCTGCTGAAG CCATTGGTGC TGATCAGCCG
 33001 ATGGGTAAGC CATTCTCCT TGGAAATCCT AAGCTCAGAA GGACCGAGTA
 33051 TCTAGTCCAT TCATGGTAAA CCATTCCAAA TAGACAGGGA GATGGGAGGG
 33101 CAAACCTGCA TTTGATTCCC AGCATCGGTT GTGCCCTTCC CTTGGTAGTA
 33151 ACAGGCTTGA TATGCAGATG GGAGCATCTC ACTGTGAGCC GGGGATTGTT
 33201 GGGAGTCTTT TTGTACCTCC CTTGCATTGG TGAATGTATT ATAGGGAAAT
 33251 AGTGAGCCAT TTTGAAATGC TTCTGAAAG GGTGAATGTC CCAGGGCATG
 33301 TGCAGAGCAA CCATCTGTTT TTGAAGATGA ATCATCTCAT GGTGGAGAGC
 33351 AGCTGTTAGC AGACACTGAG AAGCTTGTG AGTGCTCTGC GGATCAGAA
 33401 CAGCTTTCAG TCTAGGCTGG CTGATCTGCC TGGGTGTGCT TTTTATTTT
 33451 TTTTGTATTG TTTTATTTTA TTGTATTTTT TAAGACAACA GCACCTAGTA
 33501 TTTCCAGGGG CTTTCCCGTT CAAGTACGAA CCAGGCTTGA CCCTGCTTAG
 33551 CTTCCAAGAT CAGGTGAAAT TGAGCACATT CAGAAATGGTA TGGCTATAGA
 33601 CCTGGATTCT CTTTATATT TTTTATATTC TTTTTCAGTT GATTTTAACT
 33651 CGTGAGGCAT ACCAAATATA TATATGGATG CAGTATGTGT GACATTTGGA
 33701 TACATATGTA CAATGTGTAA TTATCAAATC AGGGTAATTG GCATATCCAT
 33751 CTTGTGTCTA CTTTAAATTT TCCAAATGTT TCTGCCCTTC CAAGAAGGAA
 33801 GAGGCAGGTG GTAGCTTGGT GTAACCTGTG CACTTTTCCC TGGAAAGATA
 33851 ATGGATCGGG AGCAACAGAA GCAGCCACCA TGATCCGAAG CCATAGAGGA
 33901 GAATCTGTCT TCTTCTCTAA CACCCCAAAC CCAGCTGCTG TAACTCTTCT
 33951 GCCTCCATTT GGGTATAATT TATTTGGCTA TCCTGCAGG TGTCACTCTC

FIGURE 3J

34001 CTAAGTCCAG ACTTCACAGC TCTCCAGAGG CTTTGGGGCT GCTTTGAGTT
 34051 TAATGATAGA GCCACCAGAT GATTTTTCCC AAGAGTTTTT ATTATCTATT
 34101 CATGGAGCAA GTATGACCTT TTACCAGACT CAGTCTTTAC AAGGTTGGTC
 34151 TCCTGCTTAT AGCATAAGAA CATCTTCTAG ATTTTAAATT CAACCACAGA
 34201 GAAACTCAAG GCACATATAC ACAGTCTGTA TTAGCACATT TAAATAGATT
 34251 TCCGACAAGG GAGGACAAAT GTTTCCTGCT GTTTAACACA TGAGGGTCTG
 34301 GTTTAAGGTG GAGCTTTTCT TAGGGACAGA GACCTTTCTT TTTAATGACC
 34351 AGGTGAGATC TGTAAAGTTGA TCACAGACTG TTTTCCTACT CTGTGCAGTC
 34401 AAGGCACTGG AGTAATAAAA TAGGGATATC CTGTGGTGAG TTACGTCATT
 34451 TTTGGAAGCT ACACTTGAAG CAGTAGTAGG AAGAGAGCCA TAGTGGTATG
 34501 GAAAGATGGA ATTCTGCTCT GGCCTCTTGG TCCTGCAGTG TCTTCATCTA
 34551 ATTCTAGGGA CACTGACTTG GATGGGACAG ATATAAATAG GCTGTGACA
 34601 TTTTAATTGC AATTTTGTIT TTATTTTGA AGGCATGTAC ACCTGTATGC
 34651 CCATGGCAAA GATTGAGATT TTCAAAGGT ATATAGAGAG CATTAAGCTT
 34701 CCACCCCGCG CCCTCCACTC TAGTCCCAA TTTTACAATT TCCCAATTCA
 34751 GAGGCAACCA TATCCCACTG TTCTTTTTTG TTTGTTGTTT TGTTTTGAGA
 34801 TGTTTAGTGT ATGATTGTCA TGTGGGGTGA GTGTGTGTTT TTCTCTCTC
 34851 TTTTTTCTTT TTTAAGACAA ATTGTAGCAC TCTGTAGGTA CTGTATTGCT
 34901 TCATGCTTTT TTCACCTAAA AAAAGTGATA TAAACTGTCT CCCATGATAG
 34951 TGATATGCTA TATCATGTGA TAGAGTGATA TATCATGGGG ATAGTTTCAT
 35001 ATCACACCAT CACACCTAGA GTTCTGCCTC ATACTTTGTT AAAAGCTATA
 35051 CGGGGGCACC ACGATTTACC TATCGAGTTC CCACTGGTTA ACATTTAAAT
 35101 TGTTTTCACT CTTTCTTCTT TAAATAATGC TGCAGTGAGA TATTTTGAAT
 35151 ATAAGCTTTT GTGTATGTGT GTGAGGATAT CTGTGAGGTA AATTTCTAGA
 35201 CATGAAATTG CTGGGTCCGA AGGACATGTG GGTTTGTATC CTGTAAAGT
 35251 GTCACCAAAAT CGCAATGGGA CCATTTTGCA CTCTTGCTGA TGATGTATAA
 35301 GTGTGCTGAG CAGGCTTGGA ATGTCTCCTG TCTGTTTCGG CAGGTTGTAC
 35351 CAGTCAGTTA AGCTGCTGTA CTCCATCGGG ATCTTTTCA CCTACGCACT
 35401 CCAAGTCTAC GTCCCGGCTG AGATCATCAT CCCCTTCTTT GTGTCCCGAG
 35451 CGCCCGAGCA CTGTGAGTTA GTGGTGGACC GTTTGTGCGG CACAGTGCTG
 35501 GTCTGCCTGA CATGTGAGTA GAAAGATGATA ATTGCTTGC TGTGTTTTCC
 35551 CTAAAGGGCA CCCAGTCTGC AGGCTTTTCA GAGAAAAGAC AATGTGTGTT
 35601 GTAGTGAAGC TGGCTATGTT TGTGACAGAG AACCTGGCCC ATGGCCTCAC
 35651 TTTCAAGATT GAGGCACCTC CAGATGGGGA AGTGAATTAA TTACATATGT
 35701 ACTGTAAAGA ACATGGGAAT GAGGACGTG GTTTATGTAT AGATAGGTA
 35751 TGAAATGCTG TGGAGGTGGT TATCATTGAG AGTAAAGACA TGCGATTACT
 35801 ATCCCATATT AAATAAGGTA AAGGCTGTA AGCCATTTAA CCCATATCTG
 35851 TAATGAGTAT AAGTTACTCT GATGAAGGGT ACTTATTTGC TTTTCAAAT
 35901 AGTTGTTTTT CCACTGTGAC AAGTTGCTCC TTAGATTTC TTTAGAGGCT
 35951 TTATGATAGT ATTCTAGACA TTTTAAATG TCAGTCTTAC TAAATATGTT
 36001 TCAGAAAATT TCTATTGATT AACCTAGGTA TTTGATTGAT CACTTGTGTT
 36051 TTAATCTTCT TCTCTCAACC CCAATCCGAG GAGTGTAAGT TAAAAGACAG
 36101 GATACCCCTC TGTTTGCTGT GGTGAAAAAC TGGTGACATT TAGAAAAATA
 36151 AAGTAAATTT TTTTGTAGC TTCTGTGAGT TGGTAGACTA GAGAACCCCT
 36201 GAGCAAAATCG GTTGATAATA GCTAAATTTAA GTTCTAAGA GATTTGCAAT
 36251 TGTTTTCCAA ATTCAAATGC TTTAAAGCAT AGATTCTCT TTTGGCTCT
 36301 ATTTGGCTTT TTTTCTCTT TTTAGGTTTT ATTATTTTTG AACAAGAAC
 36351 TCTTTGCTTA TTATGTTGAG ACTTCCCTGA GAATTTTCTT AAATATTCA
 36401 GTCTGAGCCT CTGTCTTGG GATAAAGATA GATCCATATG ACTTTTTAAA
 36451 TTCTAATTAG GGTGAAATGT TTTAAGGATG AAAGATGGGA AAGTTGTCTA
 36501 GCATTTGCTC TTAGTCACTC CTTCAGGCCC TCTCTAGAC CAGCCTATAT
 36551 AGAAACAGCC CACGCAGCAG CTAATCCAGG GGCCAGGGCT GTTGAAAGCC
 36601 AGCTGCTGTT CCCACAGCGA CTGAAAAAGA AGGAACATGA TGATCTCTG
 36651 TTTTCTAATA GATTGCTTTA ATGTGTGCTG CTAAGATGGG ATGCTTGGAC
 36701 TGTAAATTTT AATCCTATCT TGTGCCAGTA ACTCTCCATG CTTTGATTCC
 36751 AAAGTGTATG TTTCCACCGT GGATGGAGTA GCTCTAAGTG CTTGAGGAGA
 36801 CAGCTTTTCA GTGTATGGTA TTTATAATGT AAACCTGAGG GGCCCAATTC
 36851 TTAATCTCAA AGGGCACTGG AAGAAAGAGT GTGGTTAGTT CAATAATTT
 36901 CTCTTTATCC AAAGTGCTCC CTCCGGAAAA AGTAGGTCCT TGTAGGTA
 36951 ATGTGCTTCT CTGACTAAAC AGCTCTCTCA CCCTGCCTAT TGAGCTGGGG
 37001 CAGTGACAGG AGCCTGACTC CTCTCCCTGC CCAATTTTCC CTTCCAGCCT
 37051 GGCTCAGCCT CCCTGTAGCA TATGTACAC TTCTGCCCAG GTTTATTTCT
 37101 GCAGCACCCT GCAGGAGACA GCAGTCTCTG ATTCACAGAC CTCAATTTAT
 37151 CCTTAGATGC CTCTTGATTG TTGCTTCACT TTTCTGGGCC CTGTCTGTGA
 37201 GTCTCATCTC CTTTCAACAG GACGATGCTC AGAAGACACG GCTGCTTTTG
 37251 GTCTTCAAGT GTGTGCAAGT GTTTTTCCCT TCTGTGATC GTTGAGACT
 37301 AGCATTGCAT TGTCTCTCTG TTCAAAAAGG CAGCCCCCTT TATGTCTGAG
 37351 AGCACTCGCC TCTCTCACCT TCCTTGAGGA CTTTGAAGTA ATTGTGGGAG

FIGURE 3K

```

37401 TCAGTAGAGG CCTTTCATGG CAGCAGCAAC TTAATGTAT TTATGCGCGT
37451 TCATTTTGTG CTGCTTCTC TGTCTTTCA GATCTTTAG CACCGTTGGT
37501 TAGTATGTGA TTTTAGATCT TTAATTGATT TTTTTCATT ATATTTCATA
37551 ATTTTAAACAG CAGCTTCTTT TATTACTATT TCTGGTGTG TCCTATCTTT
37601 TCCCAACACT TTCTCTCTCC TCTTCACCT CCAAAGGGAA CAGGAGGAAT
37651 TCAGTGTAGT TCTTTTTTTT TTTTCCCTCT TGGAAATCAA CTTTCTCACC
37701 ACTCTCCCCC ATCCTCCAAA GATTACTATG GCTGATACGG ACTTTGTGAT
37751 GCTTAATTTT AAACAGTTGG AGAAGAGGGG GAGGGAACAC AAGTATTTCA
37801 TAGGATAGTG CTCATTTTGT TATGATTTCA TATCGGACAG TATCTACTTC
37851 CAGCCCATAT TTTTGGAAAT GCGGACTTAG CAGGTACCTG TATGTCCAGA
37901 CCTTGTGTGG AAGAGGCTGG CCCCACCTGT GGAGTCTGGA GTTGTAGGAT
37951 CAACGGTTTT TTAGATTTCT TTGGAGCAAT AACCCATCCA TCCTTCAGTG
38001 ATTCATACTG ATCTCTGTG TCATTTGCCA TGTGAAACAT TTTACTTCAG
38051 TTTGCTATGA AAATTTTCTG AACCTATTTT TGAAGATATA ATTACCTAAA
38101 ATCCGATCAT CCAAGAAGCC GTGCGATTTT TAGAGGATGA AGCTTCCAGG TCCAAACAGA
38151 ATTCCAAGCA GTCCGATTTT GGCCTTTTAC ACTTTTTAGT CTTCAGCTA
38201 GTAGCTTCTT AGTACCTTTG GAGCAGCATC ATTAATTAGC TGTGAAACCC TGCCACCCCC
38251 CAGTGAAGAA GAGCAGCATC GGAACCTCTA TAAGGCCTCA GGGTCTCTAA CAACTCCAAA
38301 CACCTCGCAT TCCCCGCCCA CTGACCTCTA AGGTTTCTCT GTTTCTACAC
38351 CTGTAAGATA GGAAGGGTGT TGCTTTGCGC TTAAGTTTAA GTTTGGAATA
38401 ATCTGTGATG TCTGTATAGG TGAAGTCAATG TAAAGTCAAT TTTTGTGCA
38451 AAATATTACT CTAAAAAAG AAAGTCAATG TAAAGTCAAT TTTTGTGCA
38501 AGAAGAAATT CCAAGTATTC ACCAATTTAA CAAAGTCAAT TTTTGTGCA
38551 TTGTATCTTC TGTGTCTTAA TCCTCATGGG TGCCCTGTAA AAATAGTTGC
38601 AATTGTAGTT TACACATAAT TTTGTCTTTC ACATTTTATT TAGTTTTATA
38651 TCACAAATAT TCATATCTTT CACTAAATTT TTCTGACCTT CGTGGTATTC
38701 CACTGTATTG GTGGATCATA TTAACCTAAG TACTCCTTTC ATGTTGGACA
38751 TGGTGGTTGT TTCCCTGTGT TTCGTATTTT TTAATTTTAT ACCCCCACTA
38801 AGTCAAACTT TGTATCTGT CCAAGACTAC CCCCCTCTCT ACTAAAAATA
38851 TTATAGACAT TTAAGGTAA CATGGTGAAA GGTGCTCTGA GTCCAGCTA
38901 CAAAAACAAA ATTAGCTGG TGTGGTGGCA AACCAGGAG GCAGAGCTTG
38951 CTTGGGAGGC TGAGGCAGGA GAATGGCATG AGCCTAGGTG ACAGGGCGAG
39001 CAGTGAGCCA AGATCGCGCC ACTGCACTCC AGCCTAGGTG TAAATAAAG
39051 ACTCCATCTC AAAACCAATA GCTTTGATCA TAATAAATA GCAGCAAGAG
39101 TAACTTGGTA AGTTTAAACA AATGGCCAGC GTATTTCTGA AGTTGCTTTT
39151 CTCCAGCCAC AGGAGCCATA AATGGCCAGC GTATTTCTGA AGTTGCTTTT
39201 TGTCTTTTTC AGTGCCTTTC TCTTGTGTG TATAAGTCAG CTCTTTCTGA
39251 TGCTGGTTCA AAACCAAGG CTCCAGAAATC CAGTTCCTTC TGTGAACATG
39301 ACTGTTGGCC TTATGTTGCT TCAGCAGTTT AAAAGCTCAT ATCTTTGTG
39351 TCTCTTGACT CGAAGGGAAG ATGTTTGTGA ATACTGTTG AGCCCTCTTG
39401 ACTAATCATG TGGTCGAGCT GAGGTTGTCC TCTGTCCCCC CTTTGTGACA
39451 CGCCACAGCT GAGCTGCTGC GAGGTTGTCC TTTGTGTG TGTTTATACA
39501 AATGTCTTCC TTTTGTCTG GGTGCGGTC CTGGCTGGTC GCTTCTGAG
39551 GATTAGGGGA GAGTAGGGAG AGGGCTGTTT CACTTTAATC TCACAGAAGA
39601 ATATCTACCT GTTGTAGTGT TTTATTTGCC CCACATGTCT GCATAGTCGA
39651 CATTTAGTGC CAGAAGGGGT TTTATTTGCC TTTCCATGGT GCAAAAACAGA
39701 TTGCTGCTTC TGGAGTTAGT TAAAGTCATT TTGCTGATT TGGTTTGGAC
39751 TACCCGTGCT GTTGAACCTT GGGGGCTGCT GATGCTGATT GGGAGGCTCG
39801 ATCTTCTCT TCTTCCACTT TTGTGTTAGT CAGCGGTGGT ATGTGGTTCA
39851 CCTCTGCAGT GTCACGCTTC ATGTAGGGTT TCCCTTTGCT AGCCCTGTGG
39901 GCTAGGACAG GAAGAAGGAC AAAATCTCTA TTAAGGATAT TTTTATTAGG
39951 AAGAGGAGAG AAATGTTCTT GTGAAAACAA GAACAAGTTT TTAGATTACT
40001 CATGTTTATC TTATATAGTG CGGAAGATCT TTGTTCAATC AGAGGAAAAA
40051 TATAAAATAT CATGATGAAG TTTCTGTTCC ACTTAACTCC CACCACAGAG
40101 TTCTGATCCC AATCTTCTGT TTTCTGTTCC CATGAAGTGC CCATATGTCC
40151 TGGAGCATCT CTCTGACTCC ACTTAACTAT CATGAAGTGC CTGAGGCTG
40201 TGCTAGGTCA GTATGGGAGA GGGTGGGAG ATGACAGACT CTGAGGCTG
40251 GGAAAGGCTC TGATTTGTCT CCGGAGGAG CATTGTTGGA GATGCTTGTG
40301 AAGGCTCTCT GTCTCTTGTG ACTCACACTT CTTTATCCA TACAAAACCG
40351 TGGCTGGGAG TGAAGGAATA TGGGCAAGC CATGTGTTCC AGGCTTGGAG
40401 AGTGGGTTAG AAGCTCCCTT TTGGCAAGC CATGTGTTCC AGGCTTGGAG
40451 CCCCATCGCT TTGCTGTGCC ACCCTCAGGC AGGACGTGGT GCTTCCAGT
40501 TGTCAGTGAG GTGAGGAACA TATCCAGAA CACAGTCTTA AGTGACTAAC
40551 ACTGGAGTGT ATAGTTCTCT AGAATTTTCA AGTTGGGCGA GACTTCAGAC
40601 ATCACCAGCT TACCACATTT CACAGGTAAA CGAATGAAC TGGGCTCAGA
40651 ATAGTAAGCT GATTTGCCCT GCACCACTCA GCTTGTATG GAGCAGGAG
40701 TGGTGATAAT ATTTGAAGCAT TTATTTAGT TTTTGTATG GCTGAGTTCT
40751 TTACATGAGT GATATCGTTT GACCATCTG TTTAGTAGTT GGGGAAATGA

```

FIGURE 3L

```

40801 GTCTCAAAGA GTTTAGGTAA CTAGCCACTG AGTGGTAGAG CTGAGGTTGG
40851 AGCCTGGGCA TTCAAATCC AGAGCCCTACA CCCATTCTTA TACCACCCTC
40901 TCCCTGGGGC TCAGTTCTCC CGATTGTAC CTAGTGCAG CTTCACCCCC
40951 ACAGTAACAC CTGGCCACAT CACTGCAAGG ATGACATTAG GCAGGGAGAC
41001 CCAGACCCCA GAGAGGGCAA GTGTCTTGGT GTCGGTACCA CAGCAGACCA
41051 GCTCTTTTGT TGAGCGTTAG ATGACCTATG ATGAGAATGC TGTTTTGTCA
41101 TCAGCCTACA ATTTTACCTC CAATTTTCTC CATATTCAAC CCTCAAGGTT
41151 TGGGGGATGT CACCAACTCT ATTTAGAAGC AGCTAAGGGC TGA AACATA
41201 GTAGTTTTGA GTTTCAGGGA AATTAAGGCT CTAAACCTTT TGCAACCTTC
41251 TCAAAGCCTT GAGATCGGTG AAGATGTAA TGAGAATTGC CATTATGGTT
41301 GATTAATAGA AGAAGGAAG ATGAAGAATC GCCCCAGAGT ATATAACACT
41351 GTGGCAGAGG TGGATCTGAG ACTCGATATC CTGTGACCCG TTTGCTGTCA
41401 GGCTTTGGAT CCCTTGTTCA ATTCCACTTG ACAAAGCAGA AAGAGGTGAG
41451 GGCTGATCGT GTGCTGGGTT CTCCACCATG CACCATGGTG CATCCCTGTG
41501 AAAGCTAGCC TGGGTATCTA CCTCTCATTT CCTCAGGC AGGATCCTTC
41551 CACTGATCCG CCAGACTCCC TGTCTCCCTT CCCTCTGCAT TTCCTGCCCC
41601 ATCAGGTCGT TGAGGTTATG GGCCAGGGGC TTGAGGTCCT GGATCCTGGT
41651 CCCAGCTCTG TTGCTTCTG TCGTTTACAC CCTCTGGGCC TCTCTTTCCA
41701 TAGGACTGTA GTAATATTGT GAGATTACAT ACCTGTGAAA CAGAGGCAGA
41751 TTCACTTCCA CTAGTGAGTG CTAGCAGTG CTCTCTGCTG GGTACCACTA
41801 GACATTCTGC AGTAATGGAA ATGAATGGAA ATGTGCCATG TGTGTGCTGT
41851 CCATTGCAAG AGCCACTAGG CACCACTGGT TCAATTTAAT TCAATTTCAAT
41901 CTAGTGTGTA ATGAAGAACA GGATTTAATT GCACAGCTCC AGGTAAGTG
41951 GTAAATAGCC GTCCGTGGCT ATGTTGGACA GCACAGCTCC AGGTAAGTG
42001 TGAGGCAGGA GGCATGAATC CATTCTTTCC CTGGTGTGTT AGTCCATTG
42051 CGTTGTATATA AAGAAGCACC TGAGACTGGG TAAATTTATA AGAAAAGAGG
42101 TTTATTTTGG CTGATGGCTC TGCAGGCTGT ACAGGAAGTG TGATGCCAGC
42151 ATCTGCTTCT GGTGAGGGCC TCAGGAAGCT TCTAATCATG GCAGAAGGCA
42201 AAGGGGGAGC AGGCTTTATA TGGCAAGACA GGGAGCAAGG AGAAGGGAGG
42251 TACCAGGCTC TTTTAAACAA CAGCTCTCTC AGGGAGGGCC CCAAGTCATT
42301 CATGAGGGAT TTGCCCCAC GACTCAAAAC CTTCCACCA GGCCCACTT
42351 CTGACATTGG GGATCACATT TCAACATGAA ATTTGGAGGG GATCCAAACC
42401 ATATTACCTT GTTCCACAT GTCTCTCATC TTAAGTCAAT
42451 GAGTGCTATT CTCTTTTGT TGTTTTATG ATCTGCTCTC TAAATGCAAT
42501 AGACATTCCA GTTTAAAGTG TTTCTTAAAA AGTGATCGAT GTAGACATTG
42551 CCAATCCTTC AGCTGCTCAG CCAAAGAAGC TCACCACTGT GGTGCTTCC
42601 GTTGCCTTGG ACTGAGATG TCTGGCAGTC CAGAAATGAAC TTAGAAGCAA
42651 TTAGAGTGAC TTGACTGCAT TTTCGCTTTA CCTCTCGTA GGAATCAATG AGGTAGTAGA
42701 ACCTCTCATA TAAATGTAA CAAGGCTGGG AGCATCCAGC TGATGCCAG
42751 TAAGCTCTGG ATGTCTGTAT CAAACTATAT TTGATTGCTA ACAGGTAGT
42801 CAGTAGGAAA GACAATCTGT CTGTAGCAGG ATGTACTTTT CATGCCAAA
42851 AACTAACAGG AAGTCATGCA TAACATTAA AGGTAAGACA TCCCTACTGT
42901 AAGATGAGTA CTAATGATGA CACCTGCATA ACCTCATTTG ACCATCATGA
42951 ACACCAGGCC TTTGTGAGG AGTTAATATC CCCATTTTGC CAACAAGAAA
43001 CATCTCTATG ATTCAGGAGC CCAATGTAC CCAATGTAC TCAGTCAATT
43051 ACTGGGGGAA AGAAAGGTAC CATACTTTCC CCAATGTAC GAGCCACAG
43101 AGCAGCAGAG CCAGGATCTG AACACAAGAA CCTAGTTCCA GTGGAAAGAT
43151 GCCTCAATAA ACCTGTGAAA CACTGGCCTT TGCCACCTG GGTGATTTCG
43201 CGGTGAGATG GGAAGCGTGG GGTCAAGTGG CACTAGGATG GACAATGGCT
43251 GTGAAGCCTC CTCTGCTTA CAGCACTGTC TGGCAGTGT GACAATGGCT
43301 GGTATGGCAC GGAAGCCGAT GGCACCTCCT GCGGCAGTGC ACCATTGGTC
43351 TTAGCTCAGT CTTCTTCTT GGCTCACCCG TGGCTGAGTT TCAGATGTGA
43401 GAGCCAGTGG GTGCTCTGTC ACAGAGATAC GGTGCTGCTG TGGGGCTTCG
43451 CCAGGGGTCA GCCTGCAGAT AGAACTGCTT TTTTTCACCT GTATCAAAAT
43501 GCTCTGTGAA ATGCGGTTTT ATCAGGTGT CTTTCAGAA GCGGGGGTTT
43551 CTTTCTCTAT TTGTTTCTT GTCAGTCAAG ATGTCTAAGT TTTCTAGAG
43601 GCTCCCTGAG TGGTAAGAAA ATGAGCAGCT GCTCAGGAAC GTCCACCTCT
43651 TTTTCTTCTC CCTACCCTCC CTCCTTGGGT AGAAACCAGA TTAGACCTGA
43701 ATGTAATTTG GTTCTTTTGG CACAGAAAGA TGGAAATCAT GTGCTGTAGC
43751 CGGAAAGCT GAAAGCCTGG GGACCGGAGC CAGAAGATCC GGTTCGCGT
43801 CCAAGTTCTG CTGACCTTGC AGTGGGACGC CAGTTATTCC ATGTTTCTGG
43851 GCATCTATTT CACAGAGATT GAACTGGACA ATGTCTAAGT TTTCTAGAG
43901 CTCTCAATCC TATAAGATGG ACAGATGCTG AATTCTGCTA ACAGTAGGGA
43951 GACAGACCTT TCCCAGATT C TGGGATCTT AAACAGGACT GTCTCTCTC
44001 CCTGCAGGCA TCTTGGCCAT CCTCATCCCC CGCCTGGACC TGGTCTCTC
44051 CCTGGTGGGC TCCGTGAGCA GCAGCGCCCT GGCCTCATC ATCCCAACGC
44101 TCCTGGAGGT CACCACCTTC TACTCAGAG GCATGAGCCC CCTCACCAT
44151 CTTTAAGGAC GCCCTGATCA GCATCTGGG CTCTGTGGG TTTGTGGTGG

```

FIGURE 3M

```

44201 GGACCTATGA GGCTCTCTAT GAGCTGATCC AGCCAAGCAA TGCTCCCATC
44251 TTCATCAATT CCACCTGTGC CTTATATAG GGATCTGGGT TCGTCTCTGC
44301 AGCTGCCTAC CCCTGCCCA TGTGTCCCC GTTACCTGTC CTCAGAGCCT
44351 CAGGTATGGT CCAGGCTCTG AGGAAAGTCA GGGTTGCTGT GTGGGAACCC
44401 CTCTGCCTGG CACCTGGATA CCCTGGGCCA GGTAACCTGA GGGCAGGGGA
44451 GAGGTGGGGT GGCAGACACG CAGAAGTGCT ACTAGTGACA GGGCTGCCAT
44501 CGCTCACCTG TACCTATTTA CACCAGAAC TTTCCAGCTC CCCTCATCA
44551 TGCTCTCTCC TTCTACCTG CACCCAGAAC GCTGGTGAC CTGCCCCAAC
44601 TCATTCTTAC TGCACAGTTC TGGGGCAGCG ACTTTATTTA ACAATTTTCA TGTCCCCAC
44651 CTCATGTTTT CACCTTTTAC TGGGGCAGCG ATAGATTAAG TAACGGGAA
44701 CGCCCTCTCT TTATAAAGCT GGGCTTCTTT CTCATCTCTC TCCCAATGT
44751 TGTATACTCA GTATTCTTCC TATTCGAGTC TCCAGGGGGT GGCTGGACCT
44801 ACCTGGTCA TTTGAAACAGG CCCCCAAGCT GGAGTTTTTA ATCTGGACTC
44851 TCTGGCTTGC TGTGACCCTC AAGGCAATGC TTCTCTTCCC TGGTATTCCT
44901 TAGTGTGGGT CACAGTACTG TGTCTTAGT TGCTTTAGCT CTTAAACAT
44951 ACGAAGTGTT GCCTAAACTG AAAATATTTA TCITTTATTT AAAATCAGAT
45001 TTTTGTTTTT AGACTGTCTT AGATCTGGGG CTATTACGAA TAACCTTTTC
45051 TTCAGTAAAC TTTGACTCAA CTTCTCTGCG TGAGGACTCA AAGGTTTTAA
45101 ATGTCTGCAT GGGTCTCTCG CACTCTTGGC TGAGGACTCA AAGGTTTTAA
45151 TCAGGATCGT CTAAAAATGT ATACCGAAAA GACATTGAAG GACTGCAGAA
45201 CTGTTGACCA GCCTGGTTTC GAGGGAAAGG TGGCTGAGTG AGAGGCGTAT
45251 ATGTATGGGT GCACCGGGCC CAGGCCCATG TTTGAGCTC AGCCACGCC
45301 AAAATGGGGC TGTGTGCATG CTCTGTTGCC ATCGTGCTGG GACGACCA
45351 AGGTGAAAGG ATCAGCAATG TAGCTGAGGT CAGTGTGCAC AGAGTTTGA
45401 GCTCTATTGC CACCGATGAG AGCAGCTGGT CTGACACTAC GCGCAGTGCT
45451 ATTAAGTTAA TAGACTTTAC AGAAAAGGGC AGAACCAGTG CCGGCCCA
45501 CGGTGTTGTTA CAATCAGTGG GTTGAAGG AACCACCTGA AACTAATTA
45551 CACTGCCTCT GTGGCCTGGA CTTTGAAGG CTGGGTTTCA CAAACAGCTC
45601 TGAGCCCTGT CTTTCCCCCA GAATGCCTCC AGATTGGA GCAAACCTCA
45651 TGAGGTTTGC CCTCCTCAAG GTCAGCCTTC AGATTGGA GCAAACCTCA
45701 GAGAAGGCAG AGGAAGATAC ATTGCCTTGC TGTGGGCTGC CTCTCTTCT
45751 CTCTTGGTGT GCGAAGTATT TGTGTGTGTA CGTGCGTGTG TGTGTGTGT
45801 TAGCTGTGTA TTTGTGCACG GACTCCTTTT CTCTTCTGTC CCGTCACCGA
45851 CCTGTGTAAG TAACAGACCA ACAGTTCACT CTGAAAGCTG GTTGAAGGAG
45901 GCTCTTGCTT CACTGCAGAT GGTTTTGCTT CTGTGTTTGC CCAAAGCTCA
45951 AGCAGCAAAA ATGTATCAGG IATGGCCCCA GTTTTTCTG TCTCTTCTGT
46001 TAAGGGCTGT GACCACCCA CTTCAGGTG AAGGGATGGG AAAAGTGGAC
46051 TCCAAAGCCA GGAGAGCTGA ACCTACCTAA TAATTTGTTA ACTTAGGAAT
46101 TCTCATTGTA GTGACTCCCA TTCTTCTTAA GGAGAAGGAC GATTTTCACT
46151 ATGCTATCAT TGTGACTTG TTTCTCTTAA CAGGTGCTAT GTGACCAAGA
46201 CACCTTTTCT GTTCTATGGT GGAAGCTTAA CAGGTGCTAT GTGACCAAGA
46251 ATCTAGCCGG GAGTAGCAGA GGCCTGTCT CTCTGGTTC CAAGTTCGGG
46301 AGTTACCAAA GTGGGCTCAG AAAGTGTCT CTCTGGTTC CAAGTTCGGG
46351 CTCTGGCAGC CCAGCCGCTA TCTTAGCTGT CTTTCCAGC GGTGCTAAGA
46401 GTGGTCTCAG TGAGAAGGTA GATGCCAAGT GGAGGGCCAG ACCTGTGTCC
46451 TGCCCATGTC CTCCTTGGTG GACGTTTCTG TTTACTCAGA GCTGTAGAG
46501 ACCATCCTGC CCATCCGAGT TCTGAGATTG GGACTGTGAT GTTGGGACCT
46551 GAGGACTGGA TGGTAGAATA CTGGGGTCCC CCAGCTCTTA GCAGGATGCA
46601 GGCTATTGCT TCCACACCCC TGGCCGTGAG AACGTGGTAT GTAGGAGAG (SEQ ID NO:3)

```

FEATURES:

Start: 13181
Exon: 13181-13323
Intron: 13324-17943
Exon: 17944-18034
Intron: 18035-20533
Exon: 20534-20622
Intron: 20623-21093
Exon: 21094-21189
Intron: 21190-23220
Exon: 23221-23305
Intron: 23306-23728
Exon: 23729-23947
Intron: 23948-29696
Exon: 29697-29795
Intron: 29796-32613
Exon: 32614-32780

FIGURE 3N

Intron: 32781-35343
 Exon: 35344-35513
 Intron: 35514-44007
 Exon: 44008-44277
 Stop: 44278

CHROMOSOME MAP POSITION:
 Chromosome 5

ALLELIC VARIANTS (SNPs):

DNA			
Position	Major	Minor	Domain
2064	G	A	Beyond ORF(5')
2119	A	G	Beyond ORF(5')
2121	G	T	Beyond ORF(5')
2123	G	C	Beyond ORF(5')
2125	T	G	Beyond ORF(5')
2825	G	T	Beyond ORF(5')
3288	C	T	Beyond ORF(5')
6172	A	G	Beyond ORF(5')
6462	G	T	Beyond ORF(5')
7031	C	T	Beyond ORF(5')
7671	T	-	Beyond ORF(5')
8466	-	T	Beyond ORF(5')
9097	-	T	Beyond ORF(5')
9108	-	T	Beyond ORF(5')
10170	A	G	Beyond ORF(5')
10966	A	G	Beyond ORF(5')
12987	A	C T	Beyond ORF(5')
13111	C	G	Beyond ORF(5')
13120	C	T	Beyond ORF(5')
13822	C	G	Intron
14891	G	T	Intron
15207	C	T	Intron
16162	G	T	Intron
16364	T	-	Intron
16411	T	C	Intron
16636	T	C	Intron
16802	C	T	Intron
17111	A	G	Intron
17276	T	G	Intron
17372	C	T	Intron
18317	T	G	Intron
18342	C	T	Intron
21828	C	T	Intron
22674	T	C	Intron
22683	T	C	Intron
22822	C	G	Intron
23073	G	A	Intron
23343	C	-	Intron
23396	T	A	Intron
23511	G	C	Intron
23522	C	G	Intron
23582	T	A	Intron
24977	G	A	Intron
25131	A	-	Intron
25178	G	A	Intron
25351	G	A	Intron
25380	A	G	Intron
26737	G	A	Intron
26829	G	A	Intron
27423	G	A	Intron
27735	G	A	Intron
29875	C	G	Intron
30356	G	T	Intron
31344	C	T	Intron

FIGURE 30

32570	T	G	Intron
33220	T	C	Intron
33525	T	G	Intron
34589	A	G	Intron
34832	G	T	Intron
35188	A	G	Intron
35614	G	C	Intron
37852	C	A	Intron
38643	G	A	Intron
39198	G	T	Intron
39550	T	G	Intron
42281	A	G	Intron
42321	G	A	Intron
42563	G	C	Intron
42675	G	A	Intron
42908	G	A	Intron
43358	-	G	Intron
43371	G	C	Intron
44796	G	A	Beyond ORF(3')
45820	A	G	Beyond ORF(3')

Context:

DNA
Position
2064

CTTCGACTGGACTCTGCCCATGCCCAAGATCAATGCCCTGTTCAAGTTCCTATTCCGAGT
CCCCAGCGCCCAGGAACATAGTCCTTCCAGCAGTGGCAGTAATAGGTCGCGCAGGTGGTGC
TGTGGAGCAGAGCTCCGGAGCTCAGTGAGAAAAAAGCGCGGCCGCTCAAGGGAGCACGT
GACCTCGGCCTCTGGCGTGGCGGTGGATCACGTGATGAGGTCCGGAAGCGGCTGCCGG
GCAGCAAGGAGGATGGCGAGGGGCTGATACTGAACCCGGGAAGGGTGGGCTGTGCTGAA
[G,A]
CCAGAGCCGGAGCCGGAGCTGGGGCCAGAACCCGAGCAGTGAGTTCCTCCAATGACGAGT
TCCGGCTGGCGGCGCTCGCCGCTTGGGCAGGACCCACCTCGCCTTCTCCCGGCGTGGC
AGATGCTCCAGGTCAGGCACTGGATCCGCCCGGGCTGTGGGTCCCGGACTCCTTGGCGTC
CCCCGGCCGAGCTGCGGTACGACGCTGACACCCCTCTGTGAATTGGGCGAAGCGTGGAG
ATCCCTTGTCCTCGCGCTATCTCCCTGACCTCGTGGGGTGGGATCTCACCGTCTGT

2119 GCAGTCCCCAGCGCCCAGGAACATAGTCCTTCCAGCAGTGGCAGTAATAGGTCGCGCAGGT
GGTGTGTGGAGCAGAGCTCCGGAGCTCAGTGAGAAAAAAGCGCGGCCGCTCAAGGGAG
CACGTGACCTCGGCCTCTGGCGTGGCGGTGGATCACGTGATGAGGTCCGGAAGCGGCT
GCCGGGAGCAAGGAGGATGGCGAGGGGCTGATACTGAACCCGGGAAGGGTGGGCTGTG
CTGAAGCCAGAGCCGGAGCCGGAGCTGGGGCCAGAACCCGAGCAGTGAGTTCCTCCAATG
[A,G]
CGAGTTCGGGCTGGCGGCGCTCGCCGCTTGGGCAGGACCCACCTCGCCTTCTCCCGGC
GTGGCAGATGCTCCAGGTCAGGCACTGGATCCGCCCGGGCTGTGGGTCCCGGACTCCTTG
GCGTCCCCGGGCCGAGCTGCGGTACGACGCTGACACCCCTCTGTGAATTGGGCGAAGCG
TGGAGATCCCTGTCCCTCGCGCTATCTCCCTTGACCTCGTGGGGTGGGATCTCACCGT
CCTGTTGACTGACAGGTGGGGGAACTGGGGTAGATGGTGAAGATAACCCAAAGGACCA

2121 AGTCCCCAGCGCCCAGGAACATAGTCCTTCCAGCAGTGGCAGTAATAGGTCGCCAGGTGG
TGCTGTGGAGCAGAGCTCCGGAGCTCAGTGAGAAAAAAGCGCGGCCGCTCAAGGGAGCA
CGTGACCTCGGCCTCTGGCGTGGCGGTGGATCACGTGATGAGGTCCGGAAGCGGCTGC
CGGGCAGCAAGGAGGATGGCGAGGGGCTGATACTGAACCCGGGAAGGGTGGGCTGTGCT
GAAGCCAGAGCCGGAGCCGGAGCTGGGGCCAGAACCCGAGCAGTGAGTTCCTCCAATGAC
[G,T]
AGTTCGGGCTGGCGGCGCTCGCCGCTTGGGCAGGACCCACCTCGCCTTCTCCCGGCGT
GGCAGATGCTCCAGGTCAGGCACTGGATCCGCCCGGGCTGTGGGTCCCGGACTCCTTGGC
GTCCCCGGGCCGAGCTGCGGTACGACGCTGACACCCCTCTGTGAATTGGGCGAAGCGTG
GAGATCCCTGTCCCTCGCGCTATCTCCCTTGACCTCGTGGGGTGGGATCTCACCGTCC
TGTTGACTGACAGGTGGGGGAACTGGGGTAGATGGTGAAGATAACCCAAAGGACCA

2123 TCCCCAGCGCCCAGGAACATAGTCCTTCCAGCAGTGGCAGTAATAGGTCGCCAGGTGGT
CTGTGGAGCAGAGCTCCGGAGCTCAGTGAGAAAAAAGCGCGGCCGCTCAAGGGAGCACG
TGACCTCGGCCTCTGGCGTGGCGGTGGATCACGTGATGAGGTCCGGAAGCGGCTGCC
GGCAGCAAGGAGGATGGCGAGGGGCTGATACTGAACCCGGGAAGGGTGGGCTGTGCTGA
AGCCAGAGCCGGAGCCGGAGCTGGGGCCAGAACCCGAGCAGTGAGTTCCTCCAATGACGA
[G,C]
TTCGGGCTGGCGGCGCTCGCCGCTTGGGCAGGACCCACCTCGCCTTCTCCCGGCGTGG

FIGURE 3P

CAGATGCTCCAGGTCAGGCACTGGATCCGCCCCGGCTGTGGGTCCGCGACTCCTTGGCGT
CCCCGGGGCCGAGCTGCGGTACGACGCTGACACCCCTCTGTGAATTGGGCGAAGCGTGG
GATCCCTTGTCCCTCGCGCTATCTCCCTTGACCTCGTGGGTTGGGATCTCACCGTCTG
TTTGACTGACAGGTGGGGGAACTGGGGTAGATGGTGAAGATAACCCAAAGGACCATCTA

2125 CCCAGCGCCAGGAACATAGTCCTTCCAGCAGTGGCAGTAATAGGTCGCCAGGTGGTGTCT
GTGGAGCAGAGCTCCGGAGCTCAGTGAGAAAAAGGCGCGGCGCTCAAGGGAGCACGTG
ACCTCGGCTCTGGCGTGGCGGTGGGATCAGTGATGAGGTCGGGAAGCGGTGCCGGG
CAGCAAAGGAGGATGGCGAGGGGTGATACTGAACCCGGGAAGGGTGGGTGTGCTGAAG
CCAGAGCCGGAGCGGAGCTGGGCCAGAACCCGAGCAGTGAGTTCCACTGACGAGT
[T, G]
CCGGCTGGCGGCGCTCGCGCCTTGGGCAGGACCCACTCGCCTTCTCCGGCGTGGCA
GATGCTCCAGGTCAGGCACTGGATCCGCCGGGCTGTGGGTCCGCGACTCCTTGGCGTCC
CCGGGCGGCGAGCTGCGGTACGACGCTGACACCCCTCTGTGAATTGGGCGAAGCGTGGAGA
TCCCTTGTCCCTCGCGCTATCTCCCTTGACCTCGTGGGTTGGGATCTCACCGTCTGTT
TGACTGACAGGTGGGGGAACTGGGGTAGATGGTGAAGATAACCCAAAGGACCATCTAGG

2825 AGAAGTGAAGGACGCTGGGAAGTCGTCTGGTGCAGTCCCTCCTAGGACAGTTGGAGAA
ACTGAGCCCTTACTCCGGGAAGGGTAAGGGCTGCTAAGGTCATCCAGTGAGTTAATC
GGAGACCCGGAGACTGCGACTAGAATGCAAAATGTTCTAAGCTTCAGCAGCTGTTTGTCT
TTTCGCCACACCGCTCTCTCGGGGAACTTCACTGTGAAAAGGCACTCTTTCTGTCCC
TTTCTCTTTAGTCTCTCCCTTTTAGCTGTCTGCATTTTCCACCGCTGGGTTGGATT
[G, T]
GCTCTGGGTGTGGTCCCTGTTTGTTCATTATTTTCTGAAACTCATCCTTCTGTAGGT
TTGGTTTCTAACCTTCTGCTATTCTATGTAAGTCACACCAAAATATGAAATATGAATCGG
AATGTGCTTCTGGGAAGTAGGTGGCTGAGCCGAGGTTGTGGAGAGCCCTCAGGTTAATC
TGAAGAAATGTAAGACCTTTGCTTTATTTTTTCTGTAACCTGTGAGATTGGGATTGCTT
ATTTGATGACGTTTTGACGTTATTTGAAATTTGCTGAAGATAGCATCATGGTGAATG

3288 AGAGCCCTGACGTTAACTTGAAGAATGTAAGACCTTTGCTTTATTTTTCTGTAACCTG
TCAGATTTGGGATTGCTTATTTGGATGGACGTTTTGACGTTATTTGAATTTGCTGAAGA
TAGCATCATGGTGCAATGGACAGAACAGAGATTGGGGAATCAGGATATTTGTCTAGCT
CTGCCGCTTACCTGGCAACCTTAAGTGACTCGCGTTTGGGTTTCTCAGCTAGACAGTGA
TGGAATTGAATTTAAGGGCCCTTCTGCTGTGATCTGGATGTTGTGCTATCTTTAGG
[C, T]
TTGTTTTTTGTTTGTGTTTTTAAATAGAGATGAGGTCTCACTATGCTGCCAGGCTG
ATCTCAAACCTCTGGGCTCAAGTGATCTCCACCTTGGCCTCCCAAAGTATTGCGATTA
CAGGGGTGTGAGCCAGTGCCCTGACCAAGGCTGTTTGTGTTTTTATTCCGAGAGATT
TACCCGCTGTGTACACTGAGTATCAGCCTTGCAACAAGACTTAATCTAATTTGTAGGAA
GCAGTTTCTCTGCTTATTCTCTGTGCTATAAAATCTCTCTCTTTCTCTCTATCT

6172 GCTTCACAATCCAGAGTTTTAAATGGAGCCCACTATGCAGACTTTGTTTAGTGAACCTTC
TCACTTTCTGTCCTTGAACCTCTCTGTAGTAATAATCACAATGCTATCATTAAATGAGG
TTTATTATGTTCCAGGCAACATATCTAACATTTATTTTCTCTCTATCTTCATAA
CAATATTGTGATGTAGATGTTATTAATGACATCTTCAGATGAGGAGACTGTGGCAAAG
GAGATGAATTAACCTTGTCTCAGAGTCACACTACAGACTGCAAACTCAGGTGCTTTTATT
[A, G]
TGCAGAATACCCGCTGCAGACCTAATCTGCCAGGCTCTGGGGCCAGCTTTGTTGCGA
GGGAGATTTAAGGAGGGTATATAATTAAGGTGTGGTAGAAGAACTGGACTGGGAT
GCTGGTTTGTGAGTGTCTCATCAATCTTTGATTTGACTCATCTGGGCTGGATGA
GTCAGCCTTTGTGTGGGCCCTGTTTTCTGACCCCTAAGAAGGAAGCTGGAGCTTGAC
CTTCTCTAAAGCTATACCTGGCCCTAACATTTAGTGATCTTCATGGTTGGGAGTAAAAAT

6462 GCTTTTTATTGTGCGAATACCCGCTGCAGACCTAATCTGCCAGGCTCTGGGGCCAG
CTTTGTTCCGAGGGAGATTTAAGGAGGGTATATAATTTAAGGTGTGGTAGAAGAAATAC
TGGACTGGGATGCTGGTTGCTGGACTGTCTCAATCTTTGATTGACTCATCTCTGG
GGCCTGGATGAGTCAGCCTTTGTGTGTGGGCCCTGTTTTCTGACCCCTAAGAAGGAAGC
TGGAGCTTGACCTTCTCTAAAGCTATACCTGGCCCTAACATTTAGTGATCTTCATGGTTG
[G, T]
GAGTAAAAGTGTGCGTGTGCTGTTTCAGCAGTGCTTTGTGCAAGCTGCTGAGGTC
AGCAGCTGCCCTGTAGCTGTTCTAGCATCAGACTCCTACAGGAAAAAGTCTCAATTTATG
GAATGTTCTGCTCTGGTAAAGTTGGATGGAATTTCTATCTGATGCTGTTTTAAAAACAAAT
ATGTAGAAGCCAAACCAATTTTACTTCCCTCACTGTAGACCACATAGCAACACAGTCTG
TGTCTTTGTTGATTTTTAGAAATTCATCGACAGAGGAGAAAAATACATCTGGGGAAT

7031 TCGACAGAGAGGAGAAAAATACATCTGGGGAATTTGCCGCTGCTCTGAGTTCCAAAGTCCA
AACCAATGTAATTTGTTTCAAGTAACGGATGACACTTTAGCTTGCAAACAGGGGGCCG
AATGCGTGAATCTGGTAGGAGGTAGGCTAGGTTGACCTATCATAATAAGATCATAT

FIGURE 3Q

ATTTTTGTAGTGCTTTATATAAATCTACCTATAATCAAGATTACCTAGGAAGCTAGTTA
 AAAATAAAACGCCTCTTGCTGTAATCCCATCACTTTGAGAGGCTGAGACAGGTGGATCC
 [C, T]
 TTGAGGTCAAGAGTTTGAGACCAGCCTGGCCAAACACGGAGAACTCCATCTCTACTAAAA
 ACACAAAAAATTATCTGGGCATGGTGATGGACGCTGTAATCCAGCCTCGGGAGGCT
 GAGGCAGGAGAATCGCTTGAACCCAGGAGGCGGAGGTGTCAGTGAGCAAGATCACACCA
 TTGCACTCCAGCCTGGGCAACAGAGGGAGACCATCTCAAAAAAAGAAAGACAA
 AAAGACAAAAACAACAACAAAAACATAGGCTGGGCATGGTGACTCATGCCTGTAATCC

7671 TCAAGAGTTTGACACCAGCCTGGCCAACATGATGAAACCCGCTCTCTACTAAAAATAGAA
 AAAAAATTAGCAGTTTGGGTGGCGCATGTCCTGTAATCCAGCTACTCGGGAGGCTAAGAC
 AGGAGAAATTGCTTGAACCTGGGAGGCGGAGGTGTCAGCGAGCCAAGATCGACCACTGCA
 CTCAGCCTGGGCAACAGAATGAACTCCATCTCCAGTAAATAAATTAATAATAATAA
 TAAAAATAAAATAAAATGCTAAGGTGGAATCAAGTTGGGCCAGAAATCTATTTTT
 [T, -]
 TTTCTTGACGTATGTTTCATTTAACCCAATATATCCAGATATTATCATTTGCAATATA
 TAATCAGTATAAAGATTATTAATTCATGGGATATTTCAAAATTTTTTGTTCACAGTTCA
 TTGAAATCTAGTGTGCACATTTCAATTTACCCAAGTGATTTCAAGTGTAAGATAGCTA
 TTTATGGCTAGTGGTACTGTACTGGATGGTACAACCTCAGAATATGTTACCATCTATTG
 ATCTTAATCCTCTTTATTTTGAACAAACCCAGTCACTAAAAAATTGAAATTTGGAATCCT

8466 AAAAAAATCACTGTAGAATTCCCCTTAAAAATGCCCACCTCTGAAAAATTTAACACCTAC
 AAATTTTTATTTTAAAAATAGAATAAAATTTATTTATTTTAAAAATAAAAAATTCAGT
 TTGCACATACATTTTCCATATTTGCATCCGTTGCACAAAGTGATTCACCTGCTCATTTT
 AGTGCCCATCTAAAAATGGCATTTTTGTAGATTGAAGAGCAACACTTGTCTATTTATAC
 AGCTAAACAAATAGTTACATAAGGAAAAAAGGAATGTTTAAAGTTTGTACACTTAAA
 [-, T]
 TTTTTTTTTTTTTTTTTTGGCCATCAAACCTTGACAGCTTTTTTACTCAGTTGCT
 CACTCTTCTGAGTCTAAATATCTAATGGAGATTGGACCTTGTTGCTGTTTATTTGCTCT
 CAGTAATCTGAAGGACAAAGCTTGCCTTCACTCTCACATAGTACAACCCCTCATTTAGACA
 GTTAACAGGTACTATAAATATCTCCATAGGGCGGGAACCTGGCAATTCAGCAATAGACT
 TGGCTATCAGATTTCATCAAAGGGAGCCTAAGGGCAGTGTGGCCATGGATGCCAGCACTC

9097 GCCTTGGTTTCCAAAAAGAGCCATAGAAAGAACTCCGGGGAGTGGCTCTGCCCACTGTCT
 GATGCTTGAATCCTTACATAACTGCTCTGAGAAAGGGCTTTTGCTTGGATTTTTTCAGGG
 ATAAGGGAAACAGGCTTCTCCAGAGTGATCTGTTCTATTTGGAACAGATCTGTCTTTGA
 TAGAAAGTTCTTCTTACACCTAGCAAAAAATCAGCCCTCTTGACTCTCCACGTACTGAT
 CCTAGCCCTGCTGACCTTTGAGGCCCAAAATAACAAGTCTAATCCATGTGACAGCTTTT
 [-, T]
 TTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTGTGCGCCAGGCTGGAGTGCACTGGCGC
 GATCTCGGCTCACTGCAAGCTCCGCTCCGGGTTCAACCAATCTCCTGCCTCAGCCTC
 CCGAGTAGCTGGGACTACAGGCGCCGCTACACGCGCCGGCTAATTTTTGTATTTTTGAT
 TAGAGACGGCTTTTAAAGACAGTTTTGTACCCCTCAAGTTGCTAGGTGGAACCTTCTCA
 GTGCTTCAACCAATCTCTCATTTAGTTGGTTTCTACCCCTCTTGATCTAGTCTGACG

9108 CAAAAAGAGCCATAGAAAGAACTCCGGGGAGTGGCTCTGCCCACTGTCTGATGCTTGAAT
 CCTTACATAAAGTCTCTGAGAAAGGGCTTTTGCTTGGATTTTTTCAGGGATAAGGGAACA
 GGCTTTCTCCAGAGTGATCTGTTCTATTTGGAACAGATCTGTCTTTGATAGAAAGTTCT
 TCCTTACACCTAGCAAAAAATCAGCCCTCTTGACTCTCCACGTACTGATCTAGCCCTGC
 CTGACCTTTGAGGCCCAAAATAACAAGTCTAATCCATGTGACAGCTTTTTTTTTTTTTT
 [-, T]
 TTTTGAGACGGAGTCTCGCTCTGTGCGCCAGGCTGGAGTGCACTGGCGCGATCTCGGCTC
 ACTGCAAGCTCCGCTCCGGGTTCAACCAATCTCCTGCCTCAGCCTCCGAGTAGCTG
 GGACTACAGGCGCCGCTACACGCGCCGGCTAATTTTTGTATTTTTAGTAGAGACGGCT
 TTTAAAGACAGTTTTGTACCCCTCAAGTTGCTAGGTGGAACCTTCTCAGTGCTTTCAAC
 CATTCCTCATTTAGTTGGTTTCTACCCCTCTTGATCTAGTCTGACCCCTGGATATAC

10170 TATACACAACATAAAGAGTCTCTGCAAGTGTGAGATAAATGAACATCTGTACCAA
 GTAGACAACAGAGAGTTTCTCGGTTGCTAGGGAAGGATTGGGCAATTAATAAGTCCCTG
 TATTCATCTTTCACTTCAGTAATATAGGTGTCAACCTAAAGGAAGAGTTGAGAC
 AAAAAATGCAATTTTAAACAGTTTACTTGAACCTGTTACTTGAACCAAGTGAGGACAGT
 GCGCCGGGACACATTTCAAGTTGCTGGGGAGTGCGTCTTGGCCCTTGTACACAG
 [A, G]
 TTCTTAAAGGCAAAAGCAACAGGAGAGGACTGATACAAAGTGACTTGAACGAATCT
 CATCAGTTTACAGAAATAGCATGGATTATTGATGGGCTGTACATTTTGGACTATAGGGT
 ATGAGTTATGATGCTCAGTGTAGCATTTATGACTTAGTGGTGTGAGTTAGTCTAGAAC
 CCACATAGCAAGTGGCTTCAAGAGGTAATTTAATACTCAAGGGGGAGTGACACATGAC
 TGCTCTCATTTTATGCGCTCTCTGGACCCGTAATTTAAAGGAATCTCAGATAAAAA

FIGURE 3R

- 10966 CAGGTGTGAGACACTGCACCTCAGCTGCTGTTTGCATAAATAATTATGTTTCATTGACACCT
 AGAATATTAGTGTAGAGGGAGTTGAGAGATATTTAGTTTATGCCCATGCTTTTTCGA
 CATTTGAAATGGTTTACAGGTAAGCAAACTGTTGACAGAGGTAGGCTTGGCGCCTG
 GGCCTCCTGACATACCTGTAACCTGATTACAGGCTTATACCTGTATAGCAAGAGGCTTC
 AATGCTGGTATTAAGATACTTCAGAGATTTTTTTTTTCTCCCGGCTCTAGTGAGTTTA
 [A, G]
 TTGCCCCAGAGCTGGTTGGCGCTCTTGAATTCCTCTAGCTCATGAGTAAATGAAGCTCTC
 ATAGATTTTTAGCCAAGTGGCTCTGGCAATGAAGCTAGGCAAGGATCGCTCTGGGATTTC
 CAGGTCCTTTGCTGGCATTTCGCAAGTACTTCCCTTGTGAGATAGCTTGGGGGCTCTTC
 CTACATTGCAATTGTTGAGAGAAAATGCGATCTCCGTGGATCTCTCTGGTGCCAGACTG
 GGGTGTTCCTAAAGGAGTACCTGGCACTGGACCTAAGGAGAGCCTTCGGCGGAGCACCA
- 12987 GTGTATATACGTACAG
 CAGCAGGATTGAACATTTGCACAAGGTCCAAGACATTATCTCAGAAAGGAGTAGATAATC
 CTGACCTAAGGAATAGGGAATGCGGAATCCAGGAAGCACTTCTCTTTCATTTTCCCCCA
 CTCCTCCCAAGCAGTGCCTCACTTCTGCTTGTCTAGCTGTACTCCGGAATAAAGAA
 TTTATGAGTGTAGCACCACTATACCAATGGGAAGGATGGGAGTCAGAAGTCAAGTGAAC
 [A, C, T]
 CAGCCCCCTCTGTGACTTTGCACTTTTCCATTTCCCTTGGTACCAGGCATTTTCATAC
 TTAATCCATAGTGGAGCTGTGACAGTGAGCAACTCTGACAATGACAGCTTCTACCCGAGA
 GGCACCCCAACATGGAGCTAAAGGCTCCAGCTGCAGGAGTCTTAATGCTGGCCCTGT
 CCCCAGCTGCCATGTCACGACAGAGACTTCGGAATGAAGACTACCAGACTACAGCTC
 CACGGAGCTGAGCCCTGAGGAGAGCCCTCGGAAGGCTCAACAACCTCTCTCCCCGGG
- 13111 CCTAAGGAATAGGGAATGCGGAATCCAGGAAGCACTTCTCTTTCATTTTCCCCCACTTC
 TCCCAAGCAGTGCTCACTTCTGCTTGTCTAGCTGTACTCCGGAATAAAGAAATTA
 TGAGTGTAGCACCAGTATACCAATGGGAAGGATGGGAGTCAGAAGTCAAGTGAACCTAG
 CCGGCTCTGTGACTTTGCACTTTTCCATTTCCCTTGGTACCAGGCATTTTCATCTTA
 ATCCATAGTGGAGCTGTGACAGTGAGCAACTCTGACAATGACAGCTTCTACCCAGAGGC
 [C, G]
 ACCCCAAACATGGAGCTAAAGGCTCCAGCTGCAGGAGTCTTAATGCTGGCCCTGTCCCC
 CCAGCTGCCATGTCCAGCAGAGACTTCGGAATGAAGACTACCAGACTACAGCTCCAGC
 GAGCTGAGCCCTGAGGAGAGCCCTCGGAAGGCTCAACAACCTCTCTCCCCGGGCTCC
 TACCAGGCTTTGGTCAAAGCAATAGCAACCTGAGTGTGACTTCTCTCTCTCTCTCTCT
 GGGTGGGATTCGTGTTCTAAGCTCCCTTGACTTATTTTCCCCCAATTTTCATCAGT
- 13120 TAGGGAATGCGGAATTCAGGAAGCACTTCTCTTTCATTTTCCCCCACTCTCCCAAGCA
 GTGCCTCACTTCTGCTTGTCTAGCTGTACTCCGGAATAAAGAAATTTAGAGTGTAG
 CACCAGTATACCAATGGGAAGGATGGGAGTCAGAAGTCAAGTGAACCTCAGCCGCTCT
 GTGACTTTGCACTTTTCCATTTCCCTTGGTACCAGGCATTTTCATCTTAATCCATAG
 GGAGCTGTGACAGTGAGCAACTCTGACAATGACAGCTTCTACCCAGAGGCCACCCCAAA
 [C, T]
 ATGGAGCTAAAGGCTCCAGCTGCAGGAGGTCTTAATGCTGGCCCTGTCCCCCAGCTGCC
 ATGTCCACGCAGAGACTTCGGAATGAAGACTACCAGACTACAGCTCCACGGAGCTGAGC
 CCTGAGGAGAGCCCTCGGAAGGCTCAACAACCTCTCTCCCCGGGCTCTACCAAGCGC
 TTTGGTCAAAGCAATAGCAACCTGAGTGTGACTTCTCTCTCTCTCTCTCTCTCTCTCTCT
 TCGTGTCTTAAGCCTCCCTTGACTTATTTTCCCCCAATTTTCATCAGTCTCCACTT
- 13822 TTCAGCTGTGTTGGCCCCCAAGGCTGGTGCCAAGTGAGAGCTTGGACTTAAAAAAGCT
 TCTACAGAGGACATTTCTTTTAATTTAAAGTGTGTATCTGTGCTAGAACCCCAATAAT
 TTCCAAGCATAATCGGAAGCTTCTTTGCAAAAGTCTCCCCCGAATTTTGGCCCCATCACC
 AAATCAGTATTCATTGACTGAAGAGTGGGAAGAGAGAAGAAATTAACCTTTGCACTTAA
 AAAATTCAGGTTGGTAGGAAAGGAAAGATAGACTTTGCATTCTCAAAGAGGGGCTTAAT
 [C, G]
 TCTTGTCTCCAGAACTGGGACCCAGACTCATTTGGGCTGAGTTTGGCCGCTTCAGGT
 CTCACCTTTCCCAATGTAAGAAAAATTGAGGACTCCACCAAAAGCTATGCTGGCTGT
 GTGGGGCTCACCACCTGAAATAGAAAAATCAGAGGAAGTTTGTACTCTCATTGAGTTAG
 TTTCCAGCTACTCTGATTTGAGCAGACCTCTGACTTTCTCTGTGTCCAGCAGCTCA
 GCTTTTGAGTCTGTTATTTCTCAAGCTTAGCTATTACCTTTCTGTGTTTCTGTGTG
- 14891 GAGGTCATGCTGTTATCGTAAGATACTTAAAGCTTCTTCTCTGTAGTTTCTTTGCA
 GTTTTGTCTCTTTTGTATCTCAGATCAACTTGTCTAAGCAATATTAGCAGATGAGG
 TCTGGATTTTATGTTATAGAGACATCTCTGAAGCTCAAAACCTACCACTAGCAACTT
 TAGGATAGTAGCTCATAGGTTTGGCAAAAATTATGTCTGTTTCTGTGAAATCGAACA
 AATCAGAAGATACCTTCTCAGGCTTGTATTGTGACATTTTCAGGGTATACCTTTGTTCC
 [G, T]
 AGTTTCCCTTCTGCTTGTGTTGTGATCAGTGTAGGTGACCAAGGAGCTATCTGT

FIGURE 3S

- AGTTGATGGCAGGTATTACAGTCCCATCACAGGTGGTACAAGATAAAGTAATTTGCTGGG
GCTTAGAGGACTGGTTGAGTACTTCCAGCCTGGGGCATAGGATCCACGCAAGGATTTATA
TAGAAAAACATGCCAGGTATGATTAAAGGTAGAGGTTGATTTGGAGGACCTTCTTAACCTAA
ATTAATATTTTAAATATGTCGGAAGTGTAGAGACAAGTTTTTGGCTGGGTTCTTTTAT
- 15207 CTTGATGTTGTGATACAGTGTAGGTGACCAGGGAAGCCTATCTGTAGTTGATGGCAGGTA
TTACAGTCCCACACAGGTGGTACAAGATAAAGTAATTTGCTGGGGCTTAGAGGACTTGGT
TGAGTACTTCCAGCCTGGGGCATAGGATCCACGCAAGGATTTATAGAAAAACATGCCAG
GTATGATTAAGGTAGAGGTTGATTTGGAGGACCTTCTTAACCTAAATTAATTTTAAATA
TGTGGAAGTGTAGAGACAAGTTTTTGGCTGGGTTCTTTTATATTTCTGGTTTGCCC
[C,T]
ACCCTTTTATCTAGTTTTCGCGAAGGAACAAAATACATGGAAGTACTTCTACACCTACTGC
ACATATGCATGCACACACCTGGCTCTTCTAGCAAGTCAAGGGCTCAGCAAAAACCCCTAG
TTAGGGGGTGCAATAGGAACCCCAACACTTCCATGAGTTTCTAGGGTTACTTCTTTT
ATTTTTTGGAGACAGGCTTGTCTGTGTGTCCAGGCTGGAGTGCACTGGCACAATCATG
GCTCACTGCAACCTCCATCTCTGGGCTCAAGTGATCTCCACCTAGTTTCTTAAGTA
- 16162 GCAGACTGAACTGGCTCTCTAAAGGTGAGCTGGAGTAGTCATTTGCAAAATGTGGTCTG
CACACTTTGTGGGCTTCCCAAGACATTTCAAGAAAGTCTATGAGGCTAAAACTCTCTCA
TAATAATACTAAGATGTTATCTGCTTTTCACTTGTGGATTTTGCACCTATAATGTAGA
AGCAATGGTGGGTAAATTAACACTGTAGAACGAATCAAGGCAGTGGCACCATAATTA
AGTTGTGCTTGTATTTTCACTGCCACACATGCGCAAGAAAAAGCCCTTTGCACCTAA
[G,T]
AATGTCTTTGATGAACTGTAGGATTACTAATATTTAAAAATTTGAGACCCTTCAGTATA
GGTCTTTAATATTCTGTGTGGCAAAATGGGAAGTATGCATGAAGTACTTCTATGAGTACC
AAAATATGTTACTTGTCTTAAGGCAAGACCTCGAGTGATTTATGAGTTGTCAACCAAA
CTTGCTGCCTTTTTTTTTTTCATAGAACTAGAAAGAACAACTAACCAAACTAGGTCA
TTAGAGCCCGAGTACTTGTAAAGACATTTCTTGAAATGAAAGAAATCAGCCCATCACCT
- 16364 CTGTAGAACGAATCAAGGCAGTGGCACCATAATATACTAGTTGTGCTTGTATTTTCACT
GCCACATGCGCAAGAAAAAGCCCTTGCACCTAAGAATGCTTTGATGAACTGTGA
GGATTACTAATATTTAAAAATTTGAGACCCTTCAGTATAGGTCTTTAATATTCTGTGTGG
CAAAATGGGAAGTATGCATGAAGTACTTCTATGAGTACCAAAATATGTTACTTGTCTTAA
GGCAAGACCTCGAGTGATTATAGTTGTCAACCAAACTGCTGCCTTTTTTTTTTTTT
[T,-]
CATAGAACTAGAAAGAACAACTAACAACTGTAGGTCTTTCAGACCCGAGTACTTGAAG
ACATTTTCTTGAAGAAATGAAAGAAATCAGCCCATCACTCAAGGAAACAAATAGATAATAC
ATATCTGTGCGCAGAATAAAATTCAGCTTTTCAAGCAAAATAGGAAAAAACCAACT
TGTATCCAGTACCATGAGCTTGTATAGCCCTCTACTTGAAGACTTTTCTGATGAGATTAG
TGGTGATATTAAACAAATGACTTTTTGATATTATTAATACAATGAAGATGTTAAACAT
- 16411 TGTATTTTCTACTGCCACACATGCGCAAGAAAAAGCCCTTGCACCTAAGAATGCTTT
TGATGAACTGTAGGATTACTAATATTTAAAAATTTGAGACCCTTCAGTATAGGTCTTTA
ATATTCTGTGTGGCAAAATGGGAAGTATGCATGAAGTACTTCTATGAGTACCAAAATATG
TTACTTGTCTTAAGGCAAGACCTCGAGTGATTATAGTTGTCAACCAAACTGCTGCC
CTTTTTTTTTTTTTCATAGAACTAGAAAGAACAACTAACAACTGTAGGTCTTTCAGACC
[T,C]
GAGTACTTGAAGACATTTTCTTGAAATGAAAGAAATCAGCCCATCACCTCAAGGAAAA
CAATAGATAATACATATCTGTTGCCAGAAATAAATTCAGCTTTTCAAGCAAAATAGGA
AAAAAACCAACTTGTATCCAGTACCATGAGCTTGTATAGCCCTCTACTTGAAGACTTTT
CTGATGAGATTAGTGGTGATTTAAACAAATATGACTTTTTGATATTATTAATATACAATG
AAGATGTTAACTTTGGAAGATCTGTGTAACTCAACCAAAAGTATGATGTTAGGAATTC
- 16636 AACCAAACTGCTGCCTTTTTTTTTTTTTTCATAGAACTAGAAAGAACAACTAACAACTG
TAGGTCTTTCAGACCCGAGTACTTGAAGACATTTTCTTGAAATGAAAGAAATCAGCCCT
ATCACCTCAAGGAAAAACAAATAGATAATACATATCTGTTGCCAGAAATAAATTCAGCTTT
TCAAGCAAAATAGGAAAAAACCAACTTGTATCCAGTACCATGAGCTTGTATGAGCCCT
CTACTTGAAGACTTTTCTGATGAGATTAGTGGTGATTTAAACAAATGACTTTTTTGATA
[T,C]
TATTAATATACAATGAAGATGTTAACAATTTGGAAGATCTGTGTAACTCAACCAAGTAT
GATGTTAGGAATTCGTCATGGGTAAAAGATCAATTTGAAAGAGCAAGATCACCAATGGATT
TTTTTTTTTTTTTTTTTGGAGACGTTGTCTGTGTCAACCCAGCTGGAGTGCACTGGC
ACAATCTTGGCTCACTGCAACCTCTGCTCCCGGATTCAAGCGATTCTTCTGCTCAGCC
TCCCGAGTAGCTGGGATTACAGTGCTGCCACCACTCCAGCTAATTTTTATATTTTAA
- 16802 TAAAAATCAAGCTTTCAAGCAAAATAGGAAAAAACCAACTTGTATCCAGTACCATGA
GCTTGATAGCCCTCTACTTGAAGACTTTTCTGATGAGATTAGTGGTGATTTAAACAAAT
ATGACTTTTTGATATTATTAATACAATGAAGATGTTAACTTTGGAAGATCTGTGTAA

FIGURE 3T

- ACTCAACCAAGTATGATGTTAGGAATTCGTCATGGGTAAAAGATCCATTGAAAGAGCAA
GATCACCAATGGATTTTTTTTTCTTTTTTTTGGAGACAGTCTTGCTCTGTCACCCAGG
[C, T]
TGGAGTGCAGTGGCACAATCTTGGCTCACTGCAACCTCTGCCTCCCGGATTCAAGCGATT
CTTCTGCCTCAGCTCCCGAGTAGCTGGGATTACAGGTGCCTGCCACCACTCCAGCTAA
TTTTTATATTTTAGTAGAGACGGGGTTTCGCCATGTTGGCCAGGATAGTCTCAATCTCT
TGACCTCATGATCTGCCCGCTTGGCTCCCAAAGTGCTGGGATTACAGGATGAGCCAC
TGACCTGGCTGACTTTTTTTTTTTAAATACTAAATGTATCAGGGACTCTGGCCTCTATGGT
- 17111 AGTGGCACAATCTTGGCTCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTTCTGCC
TCAGCTCCCGAGTAGCTGGGATTACAGGTGCCTGCCACCACTCCAGCTAATTTTTATA
TTTTTAGTAGAGACGGGGTTTCGCCATGTTGGCCAGGATAGTCTCAATCTCTGACCTCA
TGATCTGCCCGCTTGGCTCCCAAAGTGCTGGGATTACAGGATGAGCCACTGCACCTG
GCCTGACTTTTTTTTTTTAAATACTAAATGTATCAGGGACTCTGGCCTCTATGGT
[A, G]
TGGTGTGACTTTTATGCTGTTCACTTTGTATCTTTCTGTACAGGGTTTGGGGCTTCTGT
TATTATTATTATTATTTTAAATTCCTCTGTTCTTACCAGTGTTTGTCCGTCAITGT
TTGGTTTGTATCCTCTGTTGTCAGTTTGGGATCTGAGTCTTTTTTTTTTGTAGATTGG
AGTCTCCCTCTATTGCTAGGCTGGAGTACAGTGGCAGCATCTTAACCTCACTGCAACCTC
TGCTCCCGGTTCAAGCAATTCCTACCTTAACCTCTGAGAAGCTGGGATTACAGGC
- 17276 AATCTCTTGACCTCATGATCTGCCGCTTGGCTCCCAAAGTGCTGGGATTACAGGCAT
GAGCCACTGCACCTGGCCTGACTTTTTTTTTTTAAATACTAAATGTATCAGGGACTT
CTGGCCTCTTATGGTGTGGTGTGACTTTTATGCTGTTCACTTTGTATCTTTCTGTTACAG
GGTTTGGGGCTCTGTTATTATTATTATTATTTTAAATTCCTCTGTTCTTACCAGT
GTTTGTCCGTCAITGTTGGTTTGTATCCTCTGTTGAGTTTGGGATCTGAGTCTTTT
[T, G]
TTTTTTTGTAGATGGAGTCTCCCTCTATTGCTAGGCTGGAGTACAGTGGCAGCATCTTA
ACTCACTGCAACCTCTGCCTCCCGGTTCAAGCAATTCCTACCTTAACCTCTGAGAA
GCTGGGATTACAGGCACATGCCGCTATGCTGGCTAATTTCTGTATTTTAGTAGAGACG
GGGTTTGCCTTGTGGCCAGGCTGGTCTGCAACTCTGACCTCAGGTGATCACCCTG
CGGCTCCCAAAGTAGTGGGATTAGGCATGAGCCACTGTGCTGGCCAGGTCTGAGCC
- 17372 AAATACTAAATGTATCAGGGACTTCTGGCCTCTTATGGTGTGGTGTGACTTTTATGCTGT
TCACCTTGTATCTTTCTGTTACAGGTTTGGGGCTTCTGTTATTATTATTATTATTTT
AATTTCTCTGTTCTCTTACCAGTGTGTTGTCGTGATGTTGGTTTGTCTATCTCTGTT
GCAGTTTGGGATCTGAGTCTTTTTTTTTTTTGTAGATGGAGTCTCCCTCTATTGCTAG
GCTGGAGTACAGTGGCAGCATCTTAACCTCACTGCAACCTCTGCCTCCCGGTTCAAGCAA
[C, T]
TCTCTACCTTAACCTCTGAGAAGCTGGGATTACAGGCACATGCCGCTATGCTGGCTA
ATTTCTGTATTTTAGTAGAGACGGGGTTTGCCTTGTGGCCAGGCTGGTCTCGAACTC
CTGACCTCAGGTGATCCACCGCTTGGCTCCCAAAGTAGTGGGATTATAGGCATGAGCC
ACTGTGCTGGCCAGGTCTGAGCCTTTACAGTGGTCAAGTCAAGTGGTAGAACCCAGACCC
AAATACACTGGAAAGGATAGAGTGTCTGAAGAGAGTTGGAGCACCCTCTGGTCTAATC
- 18317 AAAAAATGCAGGCATCTGGTAAAGGTCTGCATCAGTGGAGAGGAGTGGTGACAAATTTA
GGAGGTAGCTTTTGTGTTGTTAAATGTACTGCTTTAAACATTTTAAATAGAGAAG
CATTTTAAAAAATCAGTTGACAAAAAGCGGAATTCAGACATTCATTCACCTTAAAGATAT
TTATTGAGAGTGTCTGTGCGTTAGGCACTGTTCTAAGCTCTTAGAATACATCAGTGAAT
TAAATATTCTGCCCTCATGGAGCTTACTCATGGTGGAGAGGATGACTGAGATGGCTC
[T, G]
AGCAGTTTCTGTCATAATATGAACATATGAGTTAGTTACAGATGTCTGCCCATTTTCTA
CAGTCTCCATCCGCTGTTCTTAAATGGCCAACTGCAAGAATCTTATGTTCTTTTGTG
GATTAACCTCCAGTTGACTGCCTGCCCAAAGCCATTCTGGTTCTTTTGGAGTTGAAGAG
AGACTCAGAGATGTGGGTTGCCCTTAGCTAAGTGCACTTTCTGATCTGGCATGCTGTG
TAAAGATAACTTACCCTGCTCACCTCACCTTACCTTACCTTACCTTACCTTACCTTACCTT
- 18342 TCTGCATCAGTGGAGAGGAGTGGTGACAAATTTAGGAGGTAGCTTTTGTGTTGTAA
AATGACTTGCCTTTAAACATTTTAAATAGAGAAGCATTTTAAAAAATCAGTTGACAAA
AAGCGGAATTCAGACATTCATTCACCTTAAAGATATTTATTGAGAGTGTCTGTGCGTTAG
GCACCTGTTCTAAGCTCTAGAATACATCAGTGAATTAATATTCTGCCCTCATGGAGCT
TACTTCATGGTGGAGAGGATGACTGAGATGGCTCGAGCAGTTTCTGTCAATAATATGAA
[C, T]
TAATGAGTTAGTTACAGATGTCTGCCCATTTTCTACAGTCTCCCATGCCCTGTTCTCTAAA
TGGCCAACTGCAAGAATCTTATGTCTCTTTTGTGATTTACCTCAGGTGACTGCTGCGCTAG
CCAAAGCAATTCAGTTTCTTTCGGAGTTGAAGAGAGACTCAGAGATGTGGGTTGCCCTT
AGCTAAGTGCAGTCTTTCTGATCTGGCATTGCTGAAAGATAACTTACCCGTCTCACT
CACATCCCTTAGCCGAGCTCTCCACAGTACAGGAGCCTTCTATTCTGCTGATGTGCA

FIGURE 3U

FIGURE 3V

- TTTGTCAACAAAAGAGTTAAAGTTTTATATTTTACATAGATCTACGTTCTTATTTGTAT
 TCCCATGGAAAGAGCTCGGGCATAGAGAAACCGCCACATGTTCTGTGACCCCTCTGTCC
 TAGGTACATATGATCAAACTAGCTCAGACAAATTGGTTGCTGATGATAGTCGTGAAGTT
 CTCTAAAGATTGGCTCACTGGCCACAGATTCTAAAAGCCTTGTTCAACACCTGAGCCTT
- 23396 TTTAACAAAACAGCACTTGACTCCCTAAATCTATTAAAAACAAAACAAAACACCT
 CCCCTTCTGGGAGCATTGCAATTGTATTGTAAACAGTCTTTGTATTCTTCTCTTCCACCC
 TCCAGACGTGTTGTGGACTTCTTCTGATTGTCAACAGCTGGGATTCTGCTGTGTCTAT
 TTTGTGTTTCTGGCTGACAACTTTAAACAGGTAGGCACCTGGTTAAAAAGAAAAA
 AAAAAACAGAGCGAGAATTGGCAAAAGATGATTGAAGTTTTGTATTAGGATTTTTCCA
 [T, A]
 ATCAGCTTTTGTCAACAAAAGAGTTAAAGTTTTATATTTTACATAGATCTACGTTCTT
 ATTTGATTCCCATGGAAAGAGCTCGGGCATAGAGAAACCGCCACATGTTCTGTGACCCCT
 CCTGTCTTAGGTACATATGATCAAACTAGCTCAGACAAATTGGTTGCTGATGATAGTCG
 TGAAGTTCTCTAAAGATGGCTCACTGGCCACAGATTCTAAAAGCCTTGTTCAACACCT
 GAGCCTTCTCAGGAACCTCTTCCAGCAGAGGATCCACGGCCTCTGTTGTTGAGAGG
- 23511 CCACCTCCAGACGTGTTGTGGACTTCTTCTGATTGTCAACAGCTGGGATTCTGCTGTG
 TCTATTTTGTGTTCTGGCTGACAACTTTAAACAGGTAGGCACCTGGTTAAAAAGAAAA
 AAAAAAAGAGCTCGGGCATAGAGAAACCGCCACATGTTCTGTTAGGATTTT
 TTCCAAATCAGCTTTTGTCAACAAAAGAGTTAAAGTTTTATATTTTACATAGATCTACG
 TCTTCTATTTGATTCCCATGGAAAGAGCTCGGGCATAGAGAAACCGCCACATGTTCTGTC
 [G, C]
 ACCCTCTGTCTAGGTACATATGATCAAACTAGCTCAGACAAATTGGGTTGCTGATGAT
 AGTCGTGAAGTTCTCTAAAGATGGCTCACTGGCCACAGATTCTAAAAGCCTTGTTCA
 CAGCTGAGCCTTCTCAGGAACCTCTTCCAGCAGAGGATCCACGGCCTCTGTTGTTG
 AGAGGTGTTCCGTTTCTTCTTCCCCTCATTCTAGGTGATAGAAGCGGCCAATGGGAC
 CACCAATAACTGCCACAACATGAGACGGTGATTCTGACGCCTACCATGGACTCGCGACT
- 23522 CGTGTGTGGACTTCTTCTGATTGTCAACAGCTGGGATTCTGCTGTGCTATTTTGTG
 TTTCTGGCTGACAACTTTAAACAGGTAGGCACCTGGTTAAAAAGAAAAA
 ACCAGAGCGAGAATGGCAAAAGATGATTGAAGTTTTGTTAGGATTTTCCAAATCAG
 CTTTGTGCAACAAAAGAGTTAAAGTTTTATATTTTACATAGATCTACGTTCTTATTTG
 ATTCCCATGGAAAGAGCTCGGGCATAGAGAAACCGCCACATGTTCTGTGACCCCTCTGT
 [C, G]
 CTAGGTACATATGATCAAACTAGCTCAGACAAATTGGGTTGCTGATGATAGTCGTGAAGT
 TCTCTAAAGATGGCTCACTGGCCACAGATTCTAAAAGCCTTGTTCAACACCTGAGCCT
 TTCTCAGGAACCTCTTCCAGCAGAGGATCCACGGCCTCTGTTGTTGAGAGGTGTTT
 CGTTTTCTTCTTCCCCTCATTCTAGGTGATAGAAGCGGCCAATGGGACCAACCAATACT
 GCCACAACATGAGACGGTGATTCTGACGCCTACCATGGACTCGCGACTCTACATGCTCT
- 23582 TTTCTGGCTGACAACTTTAAACAGGTAGGCACCTGGTTAAAAAGAAAAA
 ACCAGAGCGAGAATGGCAAAAGATGATTGAAGTTTTGTTAGGATTTTCCAAATCAG
 CTTTGTGCAACAAAAGAGTTAAAGTTTTATATTTTACATAGATCTACGTTCTTATTTG
 ATTCCCATGGAAAGAGCTCGGGCATAGAGAAACCGCCACATGTTCTGTGACCCCTCTGT
 CCTAGGTACATATGATCAAACTAGCTCAGACAAATTGGGTTGCTGATGATAGTCGTGAAG
 [T, A]
 TCTCTAAAGATGGCTCACTGGCCACAGATTCTAAAAGCCTTGTTCAACACCTGAGCCT
 TTCTCAGGAACCTCTTCCAGCAGAGGATCCACGGCCTCTGTTGTTGAGAGGTGTTT
 CGTTTTCTTCTTCCCCTCATTCTAGGTGATAGAAGCGGCCAATGGGACCAACCAATACT
 GCCACAACATGAGACGGTGATTCTGACGCCTACCATGGACTCGCGACTCTACATGCTCT
 CCTCTGTCCTTCTGCTGCTGCTGTTTTCATCAGGAACCTCGAGCCCTGTCCATCT
- 24977 TCCATGAGTTATGACAACCAAAATGTCTCCAGACATTGCCAAACCTTCTGGGGGGCA
 AATCGCCCCCACCAGGGGGCACTGGTTTAGACTTTTTCAATTAGATGGTTAATTCA
 TGATCATTGTATACAGTTGGAAAAATAGAGGAAAAATGTTAAGATTAAATAAAAATAAT
 TTTCTAACCTGTATTTAGATAAGTAATCTTATCAACTCCAGTTAATTTTATTGTCA
 AAATTATAAATTCACITGTTCTTGCCTCACTTAACCCATGCAGGCAAGCTGTGGGGT
 [G, A]
 GCATGAGAGAGAACATCTGTATACAGATGGGTAGAAAAATCAGGCTGAGAAAAATGTGCC
 TTAACACTATGGCTGTTTGTAAAAATGAGAATGCAATTTCTAAGGCTGAGAAAAAGGAA
 AAAAGTAAAGCGGGTAAATAAAGCATAACTTAAAAAATACTTAAATTCAGTTT
 CCCAAATAATTCATCAGTACATATTCATTAAAAATGAGACCAACCAATCTCTGAAT
 ACCATGTGCCACCCGAGTCTCTCTCAGGACCGCTGTATGTGATTGGTCTGTCTCA
- 25131 TGTTAAGATTAATAAAAAATAATTTTCTAACCTGTATTTAGATAAGTAATTCCTTAT
 CAACTCCAGTTAATTTTATTTGTCAAAATTAATAATTCAGTTGTTCTTGGCCTCACT
 AACCATGCAGGCAAGTCTGTGGGTGGCATGAGAGAGAACATCTGTATACAGATGGGTA

FIGURE 3W

GAAAAATCAGGCTGAGAAAAATGTGCCCTTAACACTATGGCTGTTTGTGAAAAATGAGAAT
 GCATTTTCTAAGGCTTGAGAAAAGGAAAAAGTAAAGCGGGTAAATAAAGCATAACTT
 [A, -]
 AAAAAAATACTTAAATTCAGTTCCTCCCAATATTCATCAGTACATATTCATTAATAAT
 GCAGACAAACAAATACCTCTTGAATACCATGTCCCCACCCGAGTCTCCTCTCAGGGAC
 CCGCTGTATGTGATTGGTCTGTCTCATTCTAGATCCTGTGAATGGATTACAGCCCATGT
 AAGTATATTGAGAAATACATTGAAATATATTTTGTTCATTTTGAACATAATTTTTT
 AAAGTTACATGTTTCATCTACCTTGCTTTTTCCACCTTAAAAATGCCTTAGTGAGCCTC

25178 AGTAATTCCTTATCACTCCAGTTAATTTTTATTTGTCAAAAATTATAAATTCACCTGTTC
 CTTGCCCTCACTTAACCCATGCAGGCAAGTCTGTGGGTGGCATGAGAGAGAACATCTGT
 ATACAGATGGGTAGAAAAATCAGGCTGAGAAAAATGTGCCCTTAAACACTATGGCTGTTG
 TGAAAAATGAGAAATGCATTTTCTAAGGCTTGAGAAAAGGAAAAAGTAAAGCGGGTAAAT
 AAAAGCATAACTTAAAAAATACTTAAATTCAGTTCCTCCCAATATTCATCAGTAC
 [G, A]
 TATTCAATTAATGACAGACAAACAAATACCTCTTGAATACCATGTCCCCACCCGAGTCT
 TCCTCTCAGGGACCCGCTGTATGTGATTTGGTCTGTCTCATTCTAGATCCTGTGAATGGAT
 TTACAGCCCATGTAAGTATATTGAGAAATACATTGAAATATATTTTGTTCATTTTTGA
 AACATAATTTTTAAAGTTACATGTTTCATCTACCTTGCTTTTTCCACCTTAAAAATGCC
 TTAGTGAGCCTCCAGGTTAGTATTCCTGGCTACCTTGTTGTTGTTCACATT

25351 CTGTTTGTGAAAAATGAGAATGCATTTTCTAAGGCTTGAGAAAAGGAAAAAGTAAAGCGG
 GGTAAATAAAGCATAACTTAAAAAATACTTAAATTCAGTTCCTCCCAATATTCATCA
 TCAGTACATATTCATTAATAATGACAGACAAACAAATACCTCTTGAATACCATGTCCCCAC
 CCGGAGTCTCCTCTCAGGGACCCGCTGTATGTGATTGGTCTGTCTCATTCTAGATCCTGT
 GAATGGATTACAGCCCATGTAAGTATATTGAGAAATACATTGAAATATATTTTGTTCCT
 [G, A]
 TTTTTGAACATAATTTTTTAAAGTTACATGTTTCATCTACCTTGCTTTTTCCACCTTAA
 AAATGCCCTTAGTGAGCCTCCAGGTTAGTATTCCTGGCTACCTTGTTGTTGTGTTGTTG
 TCACATTGTATCAGACGAAGGAGATTGCTGCCATTTATTAACAAGTCTCTCACTCAGTGT
 GCTATCAGGCCATGGATAATTTTTAGTATTTTTCAGTATTAAGACAGTGAGATGTTCTT
 ATACATTCCTTTTTGTGACTTGTATAAATACTTAAGATATTTGTCTAGATGTGAATTG

25380 AAGGCTTGAGAAAAGGAAAAAGTAAAGCGGGTAAATAAAGCATAACTTAAAAAATA
 AATCACTTAAATTCAGTTCCTCCCAATATTCATCAGTACATATTCATTAATAATGACAGACAA
 CACAAATACCTCTTGAATACCATGTCCCCACCCGAGTCTCCTCTCAGGGACCCGCTGTA
 TGTGATTGGTCTGTCTCA1CTAGATCCTGTGAATGGATTACAGCCCATGTAAGTATAT
 TGAGAAATACATTGAAATATATTTTGTTCATTTTTGAACATAATTTTTTAAAGTTAC
 [A, G]
 GTTCTCATCTACCTTGCTTTTTCCACCTTAAAAATGCCTTAGTGAGCCTCCAGGTTAGT
 ATTCTGGCTCTACCTTGTTGTTGTTAGTTGTACATTGTATCAGACGAAGGAGATTGCT
 TGCCATTTATTTAACAAGTCTCCTCAGTGGCTATCAGGCCATGGATAATTTTAGTAT
 TATTTAGTATTAAGACAGTGAGATGTTCTTATACATTCCTTTTTGTGGACTGTATATAA
 TACTTAAGATATTTGTCTAGATGTGTAATTGCTGAAGAGTGTCACCTTTGAATTTTTGT

26737 GTATCCATCCAGAGTGATGTCTATGCATAGTACAACAGGACACAGAGCAATGTCTGCAT
 AAGGGCAGCCCTGCTGATTTCTTGAGAGCAATCTGAGTCTTCTCTGGGCTTAGCCAGA
 AGTTGTGCTGTGATCAAAATAGTGCCGCTGCTGGAGTACAGCATGGGGGAAGAGGTTTG
 GCTGTGTTTTGATGTAGTCACTGCCATAGTGTGAGTTGCTTCATTTTGTGTGTCTAT
 ACAGCTAAAGATGCTCCCTTAGGTCATTTTTGTTGCCGCTGCTCTGGGCTTGTTACT
 [G, A]
 CTGTTCTGTTTTGGCATTGTGCCCACTTACCATGAGGATTCCCTACTGTTCAATGTTT
 CTGAATTTTTTCCCTAATCTAAGCATGTACATGACTGTTCTCTTGGCCCTCATGCAAG
 TGCCATTGTAGGTAGCAGACCAAGGCTCTCCACAGAGAGCAGGTTCTCTGTCTTCAG
 CATGTGGAGTCTCAATGGAAAGTCTTGGGACAGTGCTTTGCACAGAGGGTGCTCCCA
 ATAAATGTTTTTACTGTCATATCGTTGCTCTGAGATGATTTTTTATAGTTATAACA

26829 TCTGAGTCTTCTCTGGGCTTAGCCAGAAGTTGTGCTGTGATCAAAATAGTGCCGCTGCC
 TGGAGTACAGATGGGGGAAGAGGTTTGCTGTGTTTTGATGTAGTCACTGCCCATAGTG
 TTGATGTGCTTCATTTTGTGTGTCATACAGTAAAGATGCTCCCTTAGGTCATTTT
 GTTGCCGCTGCTCTGCGGCTTGTTACTACTGTTGTTTTGGCATTGTGCCCACTTAC
 CATGAGGATTCCCTACTGTTCAATGTTTCTGAATTTTTTCCCTAATCTAAGCATGTAC
 [G, A]
 TGAAGTCTTCTCTGCCCTCATGCAGTGCCATTGAGGTAGCAGACCAAGGTTCTTCCA
 CAGAGAGCAGGTTCTCTCTGTCTTCAAGTGTGAGTCTCAATGGAAAGTCTTGGGG
 AGAGTGCTTTGCACAGAGGGTGCTCCCAATAAATGTTTTATCACTGTCATATCGTTGCTC
 TGAGATGATTTTTTATAGTTATAACAGTTTCAGGATTGCAAGAGTACATCTCACAATC
 CATGTGTACCTTTAACAGCATTTTCTCAAAATACTGTTATTATAATTGATAATATGTTAA

FIGURE 3X

27423 ATGGAAGACCTCACTTAATATCATTGATACATTCTTAGAACTGCAATACATTAATGT
 ATGTATAGCGAAATCAGTTTTTTCTCATCAATGTTATAACAAAACAGCGTTGAAGGAAG
 TGACTGTACGTCACTTCACTTAAAGTCTCAGTTTCCAAGAACTATTGACGCAAGGGAG
 GACTTACTGTGTGAGAATTGAGGAGATATGTTAAACAGAGCTGATTTAAACATGTAT
 GTTTCTTATAAATAAACTTTTCTCATTAGTTGGTTGGTCAGTAGCAATCAGTAAGT
 [G,A]
 TGTAGAATAATACACTTCTCTGCTGGCCTCATTCCCACAATATCCCCACATATGGATTG
 TGAAATTCCCAGTCTGATACTTGAATCTGATCTGATGTATGAATAAGAGCAGGAGTCATT
 CACTAACCAACAGATAGCACTGTTTCCAATAACTTAGGTTACATTTGTGACTCAGGAAT
 AATTACAGGCCACTCTTGCTCTCAAGTCCCATTGTAAAGGAAAAATACCTATTACCCCTG
 CTTCAATCCAGGTATTGAAATGCTTCTTACAAGGGGATCAACAGATTTCTTAGCAGGGG

27735 ACACCTTCTCTGCTGGCCTCATTCCCACAATATCCCCACATATGGATTGTGAAATCCCA
 GTCTGATACCTTGAATCTGATCTGATGTATGAATAAGAGCAGGAGTCATTCACTAACCAAC
 AGATAGCACCTGTTTCCAATAACTTAGGTTACATTTGTGACTCAGGAATATTACAGGCC
 ACTCTTGCTCTCAAGTCCCATTGTAAAGGAAAAATACCTATTACCCCTGCTTCTTCCAG
 GTATTGAAATGCTTCTTACAAGGGGATCTAACAGATTTCTTAGCAGGGGCCAGGGAAAC
 [G,A]
 CATTTATTTAAATTTTTTATTTTTTCAAAAGCAATATTACTGCTTTGAAATCTTTCAAG
 TGAAGGCTGTTATAGAGCTTAAATATGGATCTCCTTTTACTTGCCTGAAATTTATCTGAA
 GCCTGTTAAGAGCATGCCCGTATTATCCAATAGCCATACAGTTAAATCAATTTTAAAC
 CATTTGAAAAGGCTGTTTAAACATCAATTTTTTATTTAATTGAAGCAACATACACATGTG
 GTTTAGAAAACCAATTTGAAAAAGACAGCAGCTTTGAATCCCTCCTCCCCACCTGCC

29875 GTATTAGACTTTCTTCTTAGAGAACTTGGATTGTAAAAAGGTATGACCTCTCCGAT
 TCAGAGTTCAAACTTGAATTTCTGTATAGCCTTTTGTCTTTGTTTCTGCTTTCTGCTTTT
 AGAGGATCCCAGACCCAGCCACCTCCCCTTGGTGGCCCTTGGAAAGACTACCCCTCTCT
 TCTTTGGCAGCGATTTTTTCAATTTGAAGGCAATGGAATGTTAAGAGCTGCACTGTGAT
 TTGGGCTAGTGTCTCTGGTGGCCTTGGTGTCTCAGGCTGTCTTCAAGGAATGCTGAG
 [C,G]
 AAACATTTGTTAGAAGTATCTCTGAGGCCAGGCATGGTGGCTCACGCTGTAATCTCAG
 CACTTTGGGAGGCTAGACTGGTGGATCACTTGAGGTGAGGAGTTGAAACACAGCTGGC
 CAACATGGTGAACCCCATCTCTACTAAATATACAAAATCAGCTAGGCATGGTGGCACA
 CGCCTATAATCCCAGCCACTCGAGAGGCTGAGGCAGGAGAAATGCTTGAACCTGGGAGAC
 GGAGGTTGCACTGAGCCAAGATCACGCCACTGCACTCCAGCCTGGGTGACAGAGCGAGAC

30356 CGCCTATAATCCAGCCACTCGAGAGGCTGAGGCAGGAGAAATGCTTGAACCTGGGAGAC
 GGAGGTTGCACTGAGCCAAGATCACGCCACTGCACTCCAGCCTGGGTGACAGAGCGAGAC
 TCTGTCTCAAAAAAAGAAATTTATCTTCTGTAACCTCACTGGTCACTGATGATGATGTT
 GTGAATAGTGTCTCGGGGATTCATTGAGATTTCCAGCTTCAACTTTTCAAGACAAAT
 ATATGTAATTTAAATGTTTACATTCAGGCCCTTCACTGCACACTCATCTCCTATGT
 [G,T]
 TGCAGTAAGGAATAGCATATGGCAATCAGGAAGGCAGGGTCTAGAGTCAGACTGACATGG
 GGGTAAGTCTGGCTCTGCCATAGAGTAGTGTGTGACCTTGAAGCAAGGGCTTCATCTCT
 TTGAGCCTTCATTTATTTCTGCTGAAAGTGAGCTTAATGATTCTGATTTATAGGATTA
 AATGAGATATGTCAAAATGCTTTGCAAGACCTGACACATGGTAAATGTTTAAATAGAT
 TTTTATTTTATTAATGTTATTTTATTTATTTGAATCAATAAATGCATGAATAATTTCTC

31344 GTGATCACATGTAATCACTGGTCTCCCTGAGGTTTATACCTTGGCCTGTGCTCTATCT
 TAGGGCTTTTCTTGGCCATGAAGAGTGTCTCACTGTAATTCAGAAACACAAATGGTTTC
 CTGCTGCTGGGGCAGAGGCTGAGTGGGCCATACCTCAGCAGTGAGAAAGAGATCCCAA
 GAACTCAGGACTGGAAGAAGAGGCTGAGAAAGTGGAAAGATGCCAGAGACCTTAGG
 TTCTTGGGCATCTAAGGGACCTTGTGCTAAATTTTATAGTCTTCCCTAACAGCACAG
 [C,T]
 GCAGAAATTTGTTGCTGGTTTTATACCAAGACTTGTACACAAAGTTATTCTGCAAAAC
 ATCATTTGTTTTCAAGATTTCTTGTATTTCTATTTTTTACAATAGAGAGAGAACTG
 CTAGATTGACTCTTAGTTTGGATCTAGGGCTTGTTCATTGCATCGGGGTTAAAGTGCCAG
 GCTGCACACTGTATTACCGTGTGCTCTGTGTTCTGTCAGCTGTACAGGCCAGATATGG
 GCTCCCTGCCCTCTGGCTCTTGTATTTCTTGGTATGATGGAACCTGACAGATACATTTAG

32570 GGGCCTTCTGGTGGCCTCACTCAGAGGCAAGGGACATCCAGCCTGCAGG
 GATCTCGCAGCAGGAGGACAGAGCACTGGCCTGAGCCAGGAGTCTGGGCTGCTGCTCTG
 TCCACCTTACTTTTAGGACTTCTAGCTAGGCAGTGGGCTGCAGAGTCCCTTCTAGTCC
 CAAGAGCATAACGTCTGATGAAATAACTTTATTTAAAGAGCAGATGTGCTTTTGGAGAAT
 TCTGGGGATAAAGAGTTACTTTTTTTCTGAGGTTTTTTTTTTCTTGGCCATTAACCTT
 [T,G]
 CTTTTTCTGCATTTCTCTCTCTCACTACTCTCATAGGTTCTGCCCTGGAAAA

FIGURE 3Y

- CAAAATGAAGGATCCTCGAAGTCCCACTCATCTGTACCTGGGCATGGTCATCGTCAC
 CATCTCTACATCAGCCTGGGGTGTCTGGGGTACCTGCAATTTGGAGCTAATATCCAAGG
 CAGCATAACTCCTCAACTGCCAATGCTGGTACGTGGAGGAGGATGGAACCTAGGAG
 CACTGGATATTTTAAAACTAATGGGTACAGTGTGGATTCTCCCTCTTACTTATCTCT
- 33220 CTCACTGGCTGCCCTGGACTGCATTCTGTTGGGGAATTCATGTAGAGCCTTCTGCTGAA
 GCCATTGGTGTCTGATCAGCCGATGGGTAAAGCCATTTCTCCTTGGAAATTCCTAAGCTCAGA
 AGGACCCAGTATCTAGTCCATTATGGTAAACCATTCCAAATAGACAGGGAGATGGGAGG
 GCAAACCTGCATTTGATTTCCAGCATCGTTGTGCCTCTCCCTTGGTAGTAAACAGGCTTG
 ATATGCAGATGGGAGCATCTCACTGTGAGCCGGGATGTTGGGAGTCTTTGTACCTC
 [T,C]
 CTGGCAATTGGTGAATGTATTATAGGGAATAGTGAGCCATTTTGAATGCTTCTGAAAG
 GGTGAATGTCCTCAGGGCATGTGCAGAGCAACCATCTGTTTGAAGATGAATCATCTCAT
 GGTGGAGAGCAGCTGTAGCAGACACTGAGAAGCTTGTGTAGTGTCTGCGGATCAGAAT
 CAGCTTTCAGTCTAGGCTGGCTGATCTGCCTGGGTGTGCTTTTATTTTGTATTG
 TTTTATTTTATTTGATTTTAAAGACAACAGCACTCAGTATTTCCAGGGGCTTCCCCTG
- 33525 CATTGGTGAATGTATTATAGGGAATAGTGAGCCATTTTGAATGCTTCTGAAAGGGTG
 AATGTCCCAAGGCATGTGCAGAGCAACCATCTGTTTGAAGATGAATCATCTCATGGTG
 GAGAGCAGCTGTTAGCAGACACTGAGAAGCTTGTGAGTGTCTGCGGATCAGAATCAGC
 TTTTCACTAGGCTGGCTGATCTGCCTGGGTGTGCTTTTATTTTGTATTGATTTGTTT
 ATTTTATTTGATTTTAAAGACAACAGCACTCAGTATTTCCAGGGGCTTCCCCTGCAAG
 [T,G]
 ACGAACCCAGGCTTGACCTGCTTAGCTTCCAAGATCAGGTGAATTTGAGCACATTCAGAA
 TGGTATGGCTATAGACCTGGATTGCTTTTATTTTATATTCTTTTCACTGTGATTT
 TAACCTCGTAGGACATACCAATATATATATGGATGCAGTATGTGTACATTTGGATACAT
 ATGTACAATGTGTAATTATCAAAATCAGGGAATTTGGCATATCCATCTTGTGTCTACTTT
 AAATTTCCAAATGTTTCTGCCCTTCCAAGGAAGGAGGAGGAGGTGTAGCTTGGTGAAC
- 34589 CATGAGGGTCTGGTTTAAAGTGGAGCTTGTCTAGGGACAGAGACCTTCTCTTAAATGA
 CCAGGTGAGATCTGAAGTTGATCAGACTGTTTCTACTCTGTGCAGTCAAGGCACCT
 GGAGTAATAAAATAGGATATCCTGTGGTGAATTACGTCAATTTTGGGAAGCTACACTGA
 AGCAGTAGTAGGAAGAGAGCCATAGTGGTATGGAAGATGGAATTTCTGCTCTGGCCTCTT
 GGTCTGCACTGTCTTCAATTTCTAGGGACACTGACTTGGATGGGACAGATATAAAT
 [A,G]
 GGCTTGTGACATTTTAAATGCAATTTTGTTTTATTTTGAAGCATGTACACCTGTATG
 CCGATGGCAAGATTGAGATTTTCAAAAGGTATATAGAGAGCATTAAAGCTTCCACCCCGC
 GCCCTCACTCTAGTTCCCAATTTTACAATTTCCCAATTCAGAGGCAACCATATTTCCAGT
 TTTCTTTTTTGTGTTTGTGTTTGTGATGTTTGTAGTGTATGATTGTCTCATGTGGGGTG
 AGTGTGTGTTTTTCTCTCTTTTTTCTTTTTTAAAGCAAATGTAGCACTCTGTAGGT
- 34832 CCTGCAGTGTCTTCACTAATTCTAGGGACACTGACTTGGATGGGACAGATATAAATAGG
 CTGTGACATTTTAAATGCAATTTTGTTTTATTTTGAAGCATGTACACCTGTATGCC
 CATGGCAAGATTGAGATTTTCAAAAGGTATATAGAGAGCATTAAAGCTTCCACCCCGC
 CCTCACTCTAGTTCCCAATTTTACAATTTCCCAATTCAGAGGCAACCATATTTCCAGT
 TCTTTTTTGTGTTTGTGTTTGTGATGTTTGTAGTGTATGATTGTCTCATGTGGGGTG
 [G,T]
 GTGTGTTTTTCTCTCTTTTTTCTTTTTTAAAGCAAATGTAGCACTCTGTAGGTACT
 GTATTGCTTCATGCTTTTTTCACTTAAAAAAGTGATATAAACTGTCCCATGATAGTG
 ATATGCTATATCATGTGATAGAGTATATCATGGGATAGTTTCAATACACCATCA
 CACTAGAGTTCTGCCTCATACTTTGTTAAAGCTATACGGGGGACCAAGATTACCTA
 TCGAGTCCCACTGGTTAACTTTAAATGTTTTTCACTCTTCTTAAATATGCTG
- 35188 GTACTGATGCTTCATGCTTTTTTCACTTAAAAAAGTGATATAAACTGTCCCATGA
 TAGTGATATGCTATATCATGTGATAGAGTGATATATCATGGGATAGTTTCAATACACAC
 CATCACACCTAGAGTTCTGCCTCATACTTTGTTAAAGCTATACGGGGGACCAAGATT
 ACCTATCGAGTTCCCACTGGTTAACTTTAAATGTTTTTCACTCTTCTTCTTAAATAA
 TGCTGCAGTGAGATATTTGAATATAAGCTTTTGTGTATGTGTGAGGATATCTGTGAG
 [A,G]
 TAAATTTCTAGACATGAAATGCTGGGTCCGAAGGACATGTGGGTTGTATCCTTGATAA
 GTGTCAACAAATCGCAATGGGACCATTTTGCACTCTGCTGATGATGTAAGTGTGCTG
 AGCAGGCTTGAATGTCTCTGTGTTTCCGAGGTTGTACCACTGAGTGTGAGTGTGCTG
 TACTCCATCGGATCTTTTTTCACTACGCACTCCAGTTTCACTCCGGCTGAGATCATC
 ATCCCTTCTTGTGTCGCGAGCGCCGAGCACTGTGAGTTAGTGGTGGACCTGTTGTG
- 35614 GCTTGGAAATGTCTCTGTCTGTTTCCGAGGTTGTACCACTGAGTTAAGCTGTGTACTC
 CATCGGATCTTTTTTCACTACGCACTCCAGTTTCACTCCGGCTGAGATCATCTCC
 CTCTTTGTGTCGCGAGCGCCGAGCACTGTGAGTTAGTGGTGGACCTGTTGTGCGAC

FIGURE 3Z

AGTGCTGGTCTGCTGACATGTGAGTAGAAGATGATAATTGCCTTGCTGTTTTCCCTA
 AAGGGCACCCAGTCTGCAGGCTTTCATGAGAAAAACAATGTGTGTGTAGTGAAGCTGG
 [G, C]
 TATGTTTGTGACAGAGAACCTGGCCATGGCCTCACTTTCAGAGTTGAGGCACCTCCAGA
 TGGGGAAGTGAATTAATTACATATGTACTGTAAAGAACATGGGAATGAGGACATGGTTT
 ATGTATAGATAGGGTATGAAATGCTGTGGAGGTGGTTATCATTGAGAGTAAAGACATGCG
 ATTACTATCCCATTAATAAAGGTAAGGTCTGAAAGCCATTAAACCCATGTCTGAAT
 GAGTATAAGTTACTCTGATGAAGGGTACTTATTTGCTTTTTCAAATAGTGTTTTTCCAC

37852 TTTTAAACAGCAGCTTCTTTTAACTAATTTCTGGTGTTTTCTATCTTTCCCAACTTT
 TCCTCCTCCTCTTCCACCTCCAAAGGGAACAGGAGGAATTCAGTGTAGTTTCTTTTTTT
 TTTCCCTCTTGGAAATTCACCTTCTCACCACCTCTCCCCATCTCCAAAGTACTATGG
 CTGATACGGACTTTGTGATGCTTAAATTTCAAACAGTTGGAGAAGAGGGGAGGGAACCA
 AGTATTTATAGGATAGTGCTCATTTTGTATGATTTATATCGACAGTATCTACTCC
 [C, A]
 GCCCATATTTTTGGAATGCGGACTTAGCAGGTCACTTATGTCCAGACCTTGTGTGGA
 GAGGCTGGCCCCACCTGTGGAGTCTGGAGTTGTAGGATCAACGGTTTTTTAGATTTCTTT
 GGAGCAATAACCCATCCATCCTTCAGTGATTCATACTGATCTCTGTGTCAATTTGCCATG
 TGAACATTTTACTTCAGTTTGCTATGAAATTTGAGAACTATTCTGAAGATATAAT
 TACCTAAATCGCATCATCAAGAAGCTGTTGAGACTGGAATGAGAGGCTGCAAAACAT

38643 GTCCTCAACTGTAAGATAGGAAGGGTGTCTGACCTCTAAGGTTTCTCTCAACTCCAAAT
 TCTGTGATTCTGTATAGGTGCTTTGCGCTTGATTTTAAAGTTTCTACACAAATATTACTCT
 AAAAAAGAAAGTCAATGTAATAACATTTGGGAATAAAAGAAGAAATCCAGTATTTCCAC
 CAATTTAAACAAAGTAATTTTTTTTTTGCAATTGTATCTTCTGTCTTAATCTCATGGGTG
 CCTTGAAAAATAGTTGCAATTTGAGTTTACACATAATTTTGTCTTTCACATTTTATTTA
 [G, A]
 TTTTATATCACAAATATTCATATCTTTCACTAATATTTTCATGACCTCGTGGTATTCAC
 TGTATTGGTGGATCATTAATACTAAGGTACTCTTTTCATGTTGGACATGGTGGTGTTC
 CCTGTTTTCTGATTTTTTTTAAATTTATACCCCTCAAGTCAAACCTTTGTATCTGTCCA
 AGACTACTGTGAATTTTAAAGGCATATTTATAGACATTTAAAGTAACATGGTGAACCC
 CGTCTCTACTAAAAATACAAAAACAAATTTAGCCTGGTGTGGTGGCAGGTGCTGTAGTC

39198 ATACAAAAACAAATAGCCTGGTGTGGTGGCAGGTGCCTGTAGTCCCAGCTACTTGGGA
 GGCTGAGGCAGGAGAAATGGCATGAACCCAGGAGGCAGAGCTTGCAGTGAGCCAAAGATCGC
 GCCACTGCATCTCAGCCTAGGTGACAGGGCGAGACTCCATCTCAAAACCAATAAAAAATA
 AAATAAAAAATAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA
 ATAGCAGCAAGAGCTCCAGCACAGGAGCCATAAATGGCCAGCGTATTTCTGAAGTTGCG
 [G, T]
 TTTGTTCTTTTTCAGTGCTTTGCTCTTGTGTGTATAAGTCAGCTCTTTCTGATGCTGGTT
 CAAAACCAAGGCTCCAGAAATCCAGTCTCTTGTGAACATGACTGTTGGCCTTATGTGTG
 CTTGAGCAGTTTAAAGCTCATATTTCTTGTGTCTTGTGACTCGAAGGGGAAGATGTTTTG
 TAATACTGTTGGAGCCCTCTTGACTAATCATGTGGTGCAGCTGAGGTTGCTCTGTCCC
 CCGTTTTGTACAGCCACAGCTGAGCTGCTGCTGAGAAGTGATAAAGTCATTTGTATA

39550 ATGCTGGTTCAAAACACAGGCTCCAGAATCCAGTTCCTTCTGTGAACATGACTGTTGGC
 CTTATGTTGCTTCAGCAGTTTAAAGCTCATATTTCTTGTGTCTTGTGACTCGAAGGGGA
 GATGTTTTGTAATACTGTTGGAGCCCTCTTGACTAATCATGTGGTGCAGCTGAGGTTGTC
 CTCTGCCCCCTTTTGTACAGCCACAGCTGAGCTGCTGCTGAGAAGTGATAAAGTGA
 TTTGTTATACAAATGTCTTCTTTTGTCTGGGCTGGGGTCTTGTGTGTGGGGGGG
 [T, G]
 GATTAGGGGAGAGTAGGGAGAGGGCTGTTCTGGCTGGCTGCTTCTGAGATATCTACCT
 TGTTGAGTGTCTCTCATAGGCATTTAACTCACAGAAGACATTTAGTGCCAGAAGGGGT
 TTTATTTGCCCCACATGTCTGCATAGTCGATGCTGCTTGGAGTTAGTTAAAGTCATT
 TTCCATGGTGGCAAAACAGATACCCGTGCTGTTGAACCTGGGGGCTGCTGATGCTGATT
 TGTTTGGACATCTCTCTCTTCTTCCACTTTGTGTTAGTGGGAGGCTGCTCTCTTGGC

42281 GCACAGCTCCAGGGTAAGTGTGAGGCAGGAGGCATGAATCCATTTCTTCCCTGGTGTGT
 AGTCATTTTGGCTTTGTTATAAAGAAGCAGCTGAGACTGGGTAATTTATAAAGAAAAGAGG
 TTTATTTTGGCTCATGGCTCTGAGGCTGTACAGGAAGTGTGATGCCAGGATCTGCTTCT
 GGTGAGGGCCTCAGGAAGCTTCTAATCATGGCAGAGGCAAGGGGGAGGAGGCTTTATA
 TGGCAGACAGGGAGCAAGGAGAAGGGAGGTACAGGCTCTTTTAAACAACAGCTCTCTC
 [A, G]
 GGGAGGGCCCCAAGTCATTATGAGGGATTTGCCCCACGACTCAAACCTTCCCACAG
 GCCCCACCTCTGACATTTGGGGATCACATTTCAACATGAAATTTGGAGGGGATCCAAACCA
 TATTACCTGGTAAGTCTTGTGTTTCCACATGTCTCTCATCTTACTGCAAGGAGTGTCTATTC
 TCTTTTGTGTTTTATGGCTCTCAAAATCAACTTTAGACATTTTCAGTTTAAAGTGT
 TCTTAAAAATCTGGTCTCTAAATGCAATCCATCTTCAGTCTGCTGAGCCAAAGAAGCA

FIGURE 3AA

- 42321 CATTCTTTCCCTGGTGTGTAGTCCATTTGCGTTGTTATAAGAAGCACCTGAGACTGGG
TAATTTTATAAGAGAAAGAGGTTTATTTTGGCTCATGGCTGTCAGGCTGTACAGGAAGTG
TGATGCCAGCATCTGCTTCTGGTGAGGGCCTCAGGAAGCTTCTAATCATGGCAGGAAGGCA
AAGGGGGAGCAGGCTTTATATGGCAAGACAGGAGCAAGGAGAAGGGAGGTACCGAGCTC
TTTTAAACAACAGCTCTCTCAGGGAGGGCCCCAAGTCATTATGAGGGATTGCCCCAC
[G,A]
ACTCAAACACTTCCCACAGGCCCCACCTCTGACATTGGGGATCACATTTCAACATGAA
TTTGGAGGGGATCCAAACCATATTACCTGGTAAGTCCTTGTTCACATGTCTCTCATCT
TACTGCAGGAGTGCTATTCTCTTTTGTGTGTATGGCTCCTCAAAAATCAACTTTA
GACATTTCACTTTAAAGTGTCTTAAAAATCTGGTCTCTAAATGCAATCCAATCCTTCA
GCTGCTCAGCAAGAAGCAGTGATCGATGTAGACATTGGCTGCCTTGGACTGAGATGTT
- 42563 TTAACAACAGCTCTCTCAGGGAGGGCCCCAAGTCATTATGAGGGATTGGCCCCACGA
CTCAAACACTTCCCACAGGCCCCACCTCTGACATTGGGGATCACATTTCAACATGAAAT
TTGGAGGGGATCCAAACCATATTACCTGGTAAGTCCTTGTTCACATGTCTCTCATCTT
ACTGCAGGGAGTGCTATTCTCTTTTGTGTGTATGGCTCCTCAAAAATCAACTTTAG
ACATTTCACTTTAAAGTGTCTTAAAAATCTGGTCTCTAAATGCAATCCAATCCTTCA
[G,C]
TGCTCAGCCAAAGAAGCAGTGATCGATGTAGACATTGGCTGCCTTGGACTGAGATGTTCT
GGCAGTCTACCAAGTGTGGTGCCTTCTTAGAGTGACTGACTGCATTTTTCGCTTTACAG
AATGAACCTTAGAAGCAAACTCTCATATAAAATGTAACCTCTCTGAGGAATCAATGAGG
TAGTAGATAAGCTCTGGATGTCTGTATCAAGGCTGGGAGCATCCAGCTGTAGCCAGCAG
TAGGAAAGACAATCTGTCAAACATATTTGATTGCTAACAGGTTAGTAACATACAGGAAG
- 42675 CATGAAATTTGGAGGGGATCCAAACCATATTACCTGGTAAGTCCTTGTTCACATGTCT
CTCATCTTACTGCAGGGAGTGCTATTCTCTTTTGTGTGTATGGCTCCTCAAAAATC
AATCTTAGACATTTCACTTTAAAGTGTCTTAAAAATCTGGTCTCTAAATGCAATCCAA
TCCTTCAGCTGCTCAGCCAAAGAAGCAGTGATCGATGTAGACATTGGCTGCCTTGGACTG
AGATGTTCTGGCAGTCTCACCAGTGTGGTGCCTTCTTAGAGTGACTGACTGCATTTTC
[G,A]
CTTTACAGAATGAACCTTAGAAGCAAACTCTCATATAAAATGTAACCTCTCTGAGGAAT
CAATGAGGTAGTAGATAAGCTCTGGATGTCTGTATCAAGGCTGGGAGCATCCAGCTGTAG
CCAGCAGTAGGAAGAGACAATCTGTCAAACATATTTGATTGCTAACAGGTTAGTAACAT
ACAGGAAGTCATGCACTGTAGCAGGATGTACTTTTCATGGCCAAAAGATGAGTACTAAT
GATTGATAACATTAAACAGGTAAGACATCCCTACTGTACACAGGCCCTTTTGTGAGGCACCT
- 42908 TGGACTGAGATGTTCTGGCAGTCTCACCAGTGTGGTGCCTTCTTAGAGTGACTTGACTG
CATTTTCGCTTTACAGAATGAACCTTAGAAGCAAACTCTCATATAAAATGTAACCTCTC
GTAGGAATCAATGAGGTAGTAGATAAGCTCTGGATGTCTGTATCAAGGCTGGGAGCATCC
AGCTGTAGCCAGCAGTAGGAAGACAATCTGTCAAACATATTTGATTGCTAACAGGTT
AGTAACATAACAGGAAGTCATGCACTGTAGCAGGATGTACTTTTCATGGCCAAAAGATGA
[G,A]
TACTAATGATGATAACATTAAACAGGTAAGACATCCCTACTGTACACAGGCCCTTTTGTGA
GGCACCTGCATAACCTCATTTGACCATCATGACATCTCTATGATTACAGGAGCATTAATA
TCCCCATTTTGGCAACAAGAAACTGGGGAATAGAAAGGTACCATAGCTTCCCCAATGTC
ACTCAGCTAATTAGCAGCAGAGCCAGGATCTGAACACAAGAACCTAGTTCAGAGGCCAC
AGGCCTCAATAAACCTGTGAAACACTGGCCTTGGCCACTGGTGGAAAGATCGGTGAGA
- 43358 AATAGAAAGGTACCATACCTTCCCACATGTCACTCAGCTAATTAGCAGCAGAGCCAGGAT
CTGAACACAAGAACCTAGTTCAGAGCCACAGGCCTCAATAAACCTGTGAAACACTGGC
CTTTGCCACCTGGTGAAAGATCGGTGAGATGGGAAGCGTGGGTGAGTGGGCACTAGG
ATGGGTGATTTCGGTGAAGCCTCTCTCTGCTTACAGCACTGTCTGGCAGTGTGACATG
GCTGGTATGGCAGCGAAGCCGATGGCACTCTCTGCGCAGTGACCATTTGTTCTTCGTC
[-,G]
TTCTCTCTTCTGGCTCACCCGTGGCTGAGTTTTCAGATGTGAGAGCCAGTGGGTGTCTG
TCACAGAGATACGGTCTGTGTGGGGCTTCGCCAGGGGTGAGCTGACAGATAGAACTGC
TTTTTTTCACTGTATCAAAATGCTCTGTGAAATGCGGTTTTATCAGGTGCTTTTCCAG
AAGCGGGGTTTTCTTTTCTATTGGTTTCTGTGTCAGTCAGGTAGAGATGTTTGTGTGG
AGGCTCCCTGAGTGGTAAGAAAATGAGCAGCTGCTCAGGAACGTCCACCTCCTTTTCTTC
- 43371 CATACCTTCCCACATGTCACTCAGCTAATTAGCAGCAGAGCCAGGATCTGAACACAAGAA
CCTAGTTCCAGAGCCACAGGCCTCAATAAACCTGTGAAACACTGGCCTTTGCCCACTGC
GTGGAAAGATCGGTGAGATGGGAAGCGTGGGTGAGTGGGCACTAGGATGGGTGATTTCG
GTGAAGCCTCTCTGCTTACAGCACTGTCTGGCAGTGTGACAAATGGCTGGTATGGCAG
GGAAGCCGATGGCACCTCTCTGCGCAGTGACCAATTGGTCTTCTGCACTTCTCTCTCT
[G,C]
GCTCACCCGTGGCTGAGTTTTCAGATGTGAGAGCCAGTGGGTGTCTGTACAGAGATACG

FIGURE 3BB

GTCTGTGTGTGGGGCTTCGCCAGGGGTGAGCTGCAGATAGAACTGCTTTTTTTCACCTG
TATCAAAATGCTCTGTGAAATGCCGTTTATCACGGTGTCTTTCAGAAAGCGGGGTTTC
TTTTCTATTTGGTTTCTGTGAGTCAGGTAGAGATGTTTGTGTGGAGGCTCCCTGAGT
GGTAAGAAAATGAGCAGCTGCTCAGGAACGTCCACCTCCTTTCTCTCCCTACCTCC

44796 GCCATCGCTCACCTGTACCTATTTACACCCAGAACTTCCAGCTCCCCCTCATCATGCCT
CCTCCTTCTACCTGCCTCCCTCTGCTGGTGCACCTCGCCCAACTCATTCTTACTGCAC
AGTTCACCTTTATTTAAACAATTTTATGTCCCCACCTCATGTTTTCACCTTTTACTGGGC
CAGGCATAGATTAACTAACTGGGAACGCCCTCTTTATAAAGCTGGGCTTCTTTCTCAT
CTCTCTCCCAATGTTGTATACTCAGTATTCTTCTATTGAGTCTCCAGGGGGTGGCTG
[G,A]
ACCTACCTGGTCATTTGAAACAGGCCCCCAAGCTGGAGTTTTTAATCTGGACTCTCTGGC
TTGCTGTGACCCCTAAGGCAATGCTTCTCTTCCCTGGTATTCTTAGTGTGGGTACAGT
ACTGTGTTCTTAGTGTCTTACGCTCTTAAACATACGAAGTGTGCTTAACTGAAATA
TTTATCTTTTATTTAAATCAGATTTTGTGTTTTAGACTGTCTTAGATCTGGGGCTATTA
CGAATCACTTCTTCTCAGTAACTTTGACTCAACTTCTCTGCTGAAAAGAAGCTCGCT

45820 GGGAAAAGGGCAGAACCAAGTCCCGGGCCCACTGCCTCTGTGGCTGGACTTTGAAAG
GAACCACTGAACACTAATTATGAGCCCTGTCTTCCCCCAGAATGCCTCCCTGGGTTTC
ACAAACAGCCTTGAGGTTGGCCCTCCTCAAGGTGAGCCTTCAGATTTGGGAGCAAACTTC
AGAGAAGGCAGAGGAAGATACATTGCCTTGTGTGGGCTGCCTCTTCTTCTCTTGGTG
TGCGAAGTATTTGAGAAGGCCATTGATGAATTCCTTCTTGTGATTTGTGCAC
[A,G]
TGTGTGTGTACGTGCGTGTGTGTGTGTCTGTGTAACTAACAGACCAGACTCCTTTT
CTCTTGTGTCCCGTCACCAAGGCTCTTGTCTCACTGCAGATACAGTTCACTCTGAAAGCTG
GTTGAAGGAGAGCAGCAAAAATGTATCAGGGGTTTTGTCTGTGTTTCGCCAAAGCTCA
TAAGGGCTGTGACCCACCCATATGGCCCAAGTTTTTCTGTCTCTTCTGTTCCAAAGCCA
GGAGAGCTGACTTCCAGGTGAAGGATGGGAAAAGTGGACTCTATTGTAGTGACTCCCA

FIGURE 3CC